

Are pet clinical trials ready
for prime time? p. 638

The sense (and antisense)
of ALS therapy pp. 647 & 708

Keeping families in
their homes p. 694

Science

\$15
12 AUGUST 2016
sciencemag.org

AAAS

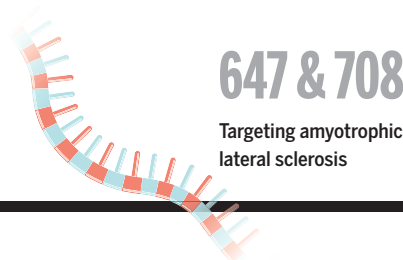


LONG IN THE TOOTH

400-year-old sharks in
Arctic waters p. 702

CONTENTS

12 AUGUST 2016 • VOLUME 353 • ISSUE 6300



NEWS

IN BRIEF

628 News at a glance

IN DEPTH

632 FOUNDER'S DEATH UNSETTLES EGYPT'S SCIENCE CITY

Nobel Prize-winning chemist Ahmed Zewail was key to fundraising and recruitment *By R. F. Service*

633 CAN CAPTIVE BREEDING SAVE MEXICO'S VAQUITA?

Scientists mull a risky strategy to save an imperiled porpoise *By B. Goldfarb*

634 NIH PLANS TO FUND HUMAN-ANIMAL CHIMERA RESEARCH

Agency would subject experiments—such as growing human organs in pigs—to extra ethical review *By J. Kaiser*

635 PARTICLE NO-SHOW AT LHC PROMPTS ANXIETY

Accelerator enters prime discovery phase, but the clock is ticking *By A. Cho*

637 SCIENTISTS SPLIT ON OREGON OLD-GROWTH FOREST PLAN

The northern spotted owl is again the center of controversy *By W. Cornwall*

FEATURES

638 FROM BARK TO BEDSIDE

Clinical trials in pets are ramping up, but will they benefit human health? *By D. Grimm*

641 OIL IN THE FOREST

A rash of spills in the Amazon has scientists and villagers scrambling to understand oil's effects on the ecosystem *By B. Fraser*

INSIGHTS

PERSPECTIVES

644 MAKING SENSE OF THE EXOPLANET ZOO

The diversity of substellar objects may call for a rethinking of our labeling conventions

By R. Oppenheimer

► RESEARCH ARTICLE P. 673

646 TWISTS AND TURNS IN GATING ION CHANNELS WITH VOLTAGE

The EAG potassium channel may use an alternative voltage-dependent gating mechanism *By G. E. S. Toombes and K. J. Swartz*

► RESEARCH ARTICLE P. 664

647 ONE TARGET FOR AMYOTROPHIC LATERAL SCLEROSIS THERAPY?

Targeting a single protein reduces both toxic repeat RNAs and proteins *By S. Mizielinska and A. M. Isaacs*

► REPORT P. 708

648 MEASURING THE PASSAGE OF BRAIN TIME

Many brain areas are under both circadian and homeostatic control *By C. A. Czeisler*

► REPORT P. 687

650 ATMOSPHERIC RADICAL CHEMISTRY REVISITED

Sunlight may directly drive previously unknown organic reactions at environmental surfaces *By V. Vaida*

► REPORT P. 699

651 BALANCING ECONOMIC AND ECOLOGICAL GOALS

What are the trade-offs between economic development and ecosystem conservation?

By E. G. Frank and W. Schlenker

POLICY FORUM

653 THE CHALLENGE OF CLIMATE-CHANGE NEOSKEPTICISM

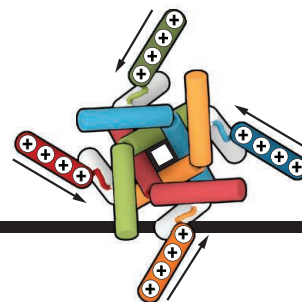
Decision science and risk management are underutilized *By P. C. Stern et al.*



Science Staff	626
New Products	713
Science Careers	714

CONTENTS

12 AUGUST 2016 • VOLUME 353 • ISSUE 6300



646 & 664

A variation on
channel gating

BOOKS ET AL.

655 DANGEROUS LIAISONS

An insider's guide to venomous creatures delivers on the science but is short on tales of the bitten and stung *By C. Kemp*

656 IN REAL LIFE

An essayist offers an intimate glimpse into the digital technologies that are redefining the human experience *By J. Golbeck*

LETTERS

657 CHINA'S ECOSYSTEMS: FOCUS ON BIODIVERSITY

By R. Wu

657 CHINA'S ECOSYSTEMS: OVERLOOKED SPECIES

By Z. Jiang et al.

657 CHINA'S ECOSYSTEMS: SACRIFICING THE POOR

By J. A. Zinda et al.

658 TECHNICAL COMMENT ABSTRACTS

RESEARCH

IN BRIEF

660 From *Science* and other journals

REVIEW

663 ZIKA

Assessing the global threat from Zika virus *J. Lessler et al.*

REVIEW SUMMARY; FOR FULL TEXT:

dx.doi.org/10.1126/science.aaf8160

RESEARCH ARTICLES

664 STRUCTURAL BIOLOGY

Structure of the voltage-gated K⁺ channel Eag1 reveals an alternative voltage sensing mechanism *J. R. Whicher and R. MacKinnon*

► PERSPECTIVE P. 646

669 PHYSICS

Laser spectroscopy of muonic deuterium *R. Pohl et al.*

673 EXTRASOLAR PLANETS

Direct imaging discovery of a Jovian exoplanet within a triple-star system *K. Wagner et al.*

► PERSPECTIVE P. 644



644 & 673

REPORTS

678 ORGANOMETALLICS

Isolation and structural and electronic characterization of salts of the decamethylferrocene dication *M. Malischewski et al.*

682 SOFT ELECTRONICS

Highly stretchable, transparent ionic touch panel *C.-C. Kim et al.*

687 SLEEP RESEARCH

Local modulation of human brain responses by circadian rhythmicity and sleep debt *V. Muto et al.*

► PERSPECTIVE P. 648

691 BRAIN MICROCIRCUITS

Imprinting and recalling cortical ensembles *L. Carrillo-Reid et al.*

694 ECONOMIC POLICY

The impact of homelessness prevention programs on homelessness *W. N. Evans et al.*

699 ATMOSPHERIC CHEMISTRY

Atmospheric photochemistry at a fatty acid-coated air-water interface *S. Rossignol et al.*

► PERSPECTIVE P. 650

702 LIFE HISTORY

Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (*Somniosus microcephalus*) *J. Nielsen et al.*

► PODCAST

705 HYDROLOGY

Continental patterns of submarine groundwater discharge reveal coastal vulnerabilities *A. H. Sawyer et al.*

708 NEURODEGENERATION

Spt4 selectively regulates the expression of *C9orf72* sense and antisense mutant transcripts *N. J. Kramer et al.*

► PERSPECTIVE P. 647

DEPARTMENTS

627 EDITORIAL

EIT's teething problems
By Helga Nowotny and Jana Kolar

718 WORKING LIFE

The 1-hour workday
By Jeffrey J. McDonnell

ON THE COVER



A Greenland shark (*Somniosus microcephalus*). These creatures are among the world's largest carnivorous sharks and are distributed throughout Arctic waters. For decades, the longevity of this slow-growing species has remained a mystery. Using well-established radiocarbon dating techniques applied on Greenland shark eye lens tissue, Nielsen *et al.* estimate that these animals are the longest-living vertebrates known to date. See page 702. Photo: © Nick Caloyianis/National Geographic Creative

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2016 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): \$165 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$1522. Foreign postage extra: Mexico, Caribbean (surface mail) \$55; other countries (air assist delivery) \$89. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #R1254 88122. Publications Mail Agreement Number 1069624. Printed in the U.S.A. Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$15.00 current issue, \$20.00 back issue prepaid includes surface postage; bulk rates on request. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that \$35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

Editor-in-Chief Jeremy Berg

Executive Editor Monica M. Bradford **News Editor** Tim Appenzeller

Deputy Editors Lisa D. Chong, Andrew M. Sugden(UK), Valda J. Vinson, Jake S. Yeston

Research and Insights

DEPUTY EDITOR, EMERITUS Barbara R. Jasny **SR. EDITORS** Caroline Ash(UK), Gilbert J. Chin, Julia Fahrenkamp-Uppenbrink(UK), Pamela J. Hines, Stella M. Hurlley(UK), Paula A. Kiberstis, Marc S. Lavine(Canada), Kristen L. Mueller, Ian S. Osborne(UK), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Jelena Stajic, Peter Stern(UK), Phillip D. Szuroni, Sacha Vignieri, Brad Wible, Nicholas S. Wigginton, Laura M. Zahn **ASSOCIATE EDITORS** Brent Grocholski, Keith T. Smith **ASSOCIATE BOOK REVIEW EDITOR** Valerie B. Thompson **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Jeffrey E. Cook, Chris Filatreau, Cynthia Howe, Barbara P. Ordway, Catherine Wolner **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Anita Wynn **PUBLICATIONS ASSISTANTS** Nida Masiulis, Dona Mathieu, Le-Toya Mayne Flood, Shannon McMahon, Scott Miller, Jerry Richardson, Alice Whaley(UK), Brian White **EXECUTIVE ASSISTANT** Anna Bashkirova **ADMINISTRATIVE SUPPORT** Janet Clements(UK), Lizanne Newton(UK)

News

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Richard Stone **DEPUTY NEWS EDITORS** Elizabeth Culotta, David Grimm, Eric Hand David Malakoff, Leslie Roberts **CONTRIBUTING EDITOR** Martin Enserink(Europe) **SR. CORRESPONDENTS** Daniel Clery(UK), Jeffrey Mervis, Elizabeth Pennisi **NEWS WRITERS** Adrian Cho, Jon Cohen, Jennifer Couzin-Frankel, Carolyn Gramling, Jocelyn Kaiser, Catherine Matacic, Kelly Servick, Robert F. Service, Erik Stokstad(Cambridge, UK) **INTERNS** Jessica Boddy, Ben Panko **CONTRIBUTING CORRESPONDENTS** John Bohannon, Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt(Berlin), Andrew Lawler, Mitch Leslie, Charles C. Mann, Eliot Marshall, Virginia Morell, Dennis Normile(Shanghai), Heather Pringle, Tania Rabesandratna(London), Emily Underwood, Gretchen Vogel(Berlin), Lizzie Wade(Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein(Edinburgh), Maggie Kuo **COPY EDITORS** Julia Cole, Dorie Cheylen, Jennifer Levin (Chief) **ADMINISTRATIVE SUPPORT** Jessica Adams

Executive Publisher Rush D. Holt

Publisher Bill Moran **Chief Digital Media Officer** Rob Covey

BUSINESS OPERATIONS AND PORTFOLIO MANAGEMENT DIRECTOR Sarah Whalen **PRODUCT DEVELOPMENT DIRECTOR** Will Schweitzer **PRODUCT DEVELOPMENT ASSOCIATE** Hannah Heckner **BUSINESS SYSTEMS AND FINANCIAL ANALYSIS DIRECTOR** Randy Yi **MANAGER OF FULFILLMENT SYSTEMS** Neal Hawkins **SYSTEMS ANALYST** Nicole Mehmedovich **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman **BUSINESS ANALYSTS** David Garrison, Michael Hardesty Meron Kebede, Sandy Kim **FINANCIAL ANALYST** Drew Shier **DIRECTOR, COPYRIGHTS LICENSING SPECIAL PROJECTS** Emilie David **PERMISSIONS ASSOCIATE** Elizabeth Sandler **RIGHTS, CONTRACTS, AND LICENSING ASSOCIATE** Lili Kiser **RIGHTS & PERMISSIONS ASSISTANT** Alexander Lee

MARKETING DIRECTOR Elise Swinehart **ASSOCIATE MARKETING DIRECTOR** Stacey Burke Bowers **MARKETING ASSOCIATE** Steven Goodman **CREATIVE DIRECTOR** Scott Rodgers **SENIOR ART ASSOCIATES** Paula Fry **ART ASSOCIATE** Kim Huynh

FULFILLMENT SYSTEMS AND OPERATIONS membership@aaas.org **MANAGER, MEMBER SERVICES** Pat Butler **SPECIALISTS** Terrance Morrison, Latasha Russell **MANAGER, DATA ENTRY** Mickie Napoleoni **DATA ENTRY SPECIALISTS** Brenden Aquilino, Fiona Giblin **MARKETING ASSOCIATE** Isa Sesay-Bah

PUBLISHER RELATIONS, EASTERN REGION Keith Layson **PUBLISHER RELATIONS, WESTERN REGION** Ryan Rexroth **SALES RESEARCH COORDINATOR** Aiesha Marshall **MANAGER, SITE LICENSE OPERATIONS** Iqoo Edim **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT ANALYST** Judy Lillibridge

WEB TECHNOLOGIES **PORTFOLIO MANAGER** Trista Smith **TECHNICAL MANAGER** Chris Coleman **PROJECT MANAGER** Nick Fletcher **DEVELOPERS** Ryan Jensen, Jimmy Marks, Brandon Morrison **BUSINESS ANALYST** Christina Wofford

DIGITAL MEDIA DIRECTOR OF ANALYTICS Enrique Gonzales **DIGITAL REPORTING ANALYST** Eric Hossinger **SR. WEB PRODUCER** Sarah Crespi **WEB PRODUCER** Alison Crawford **VIDEO PRODUCER** Nguyen Nguyen **SOCIAL MEDIA PRODUCER** Brice Russ

DIRECTOR OF OPERATIONS PRINT AND ONLINE Elizabeth Harman **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegel **PROJECT ACCOUNT MANAGER** Tara Kelly **DIGITAL PRODUCTION MANAGER** Lisa Stanford **ASSISTANT MANAGER** **DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Antoinette Hodal, Lori Murphy, Anthony Rosen **CONTENT SPECIALISTS** Jacob Hedrick, Kimberley Oster

DESIGN DIRECTOR Beth Rakouskas **DESIGN EDITOR** Marcy Atarod **SENIOR DESIGNERS** Garvin Grullón, Chrystal Smith **GRAPHICS MANAGING EDITOR** Alberto Cuadra **SENIOR SCIENTIFIC ILLUSTRATORS** Chris Bickel, Katharine Suttiff **SCIENTIFIC ILLUSTRATOR** Valerie Altonoun **INTERACTIVE GRAPHICS EDITOR** Jia You **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **SENIOR PHOTO EDITOR** Christy Steele **PHOTO EDITOR** Emily Petersen

DIRECTOR, GLOBAL COLLABORATION, CUSTOM PUBLICATIONS, ADVERTISING Bill Moran **EDITOR, CUSTOM PUBLISHING** Sean Sanders: 202-326-6430 **ADVERTISING MARKETING MANAGER** Justin Sawyers: 202-326-7061 **science_advertising@aaas.org** **ADVERTISING SUPPORT MANAGER** Karen Foote: 202-326-6740 **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **SR. PRODUCTION SPECIALIST/GRAPHIC DESIGNER** Amy Hardcastle **SR. TRAFFIC ASSOCIATE** Christine Hall **SALES COORDINATOR** Shirley Young **ASSOCIATE DIRECTOR, COLLABORATION, CUSTOM PUBLICATIONS/CHINA/TAIWAN/KOREA/SINGAPORE** Ruolei Wu: +86-186 0082 9345, rwu@aaas.org **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu + 81-532-81-5142 asandhu@aaas.org **EAST COAST/E. CANADA** Laurie Faraday: 508-747-9395, FAX 617-507-8189 **WEST COAST/W. CANADA** Lynne Stickrod: 415-931-9782, FAX 415-520-6940 **MIDWEST** Jeffrey Dembski: 847-498-4520 x3005, Steven Loerch: 847-498-4520 x3006 **UK EUROPE/ASIA** Roger Gonçalves: TEL/FAX +41 43 243 1358 **JAPAN** Katsuyoshi Fukamizu(Tokyo): +81-3-3219-5777 kfukamizu@aaas.org **CHINA/TAIWAN** Ruolei Wu: +86-186 0082 9345, rwu@aaas.org

WORLDWIDE ASSOCIATE DIRECTOR OF SCIENCE CAREERS Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532 tholmes@science-int.co.uk **CLASSIFIED advertise@sciencecareers.org** **U.S. SALES** Tina Burks: 202-326-6577 **Nancy Toerna**: 202-326-6578 **EUROPE/ROW SALES** Sarah Lelarge **SALES ASSISTANT** Kelly Grace **JAPAN** Hiroyuki Mashiki(Kyoto): +81-75-823-1109 hmashiki@aaas.org **CHINA/TAIWAN** Ruolei Wu: +86-186 0082 9345 rwu@aaas.org **MARKETING MANAGER** Allison Pritchard **MARKETING ASSOCIATE** Airnee Aponte

AAAS BOARD OF DIRECTORS, CHAIR Geraldine L. Richmond **PRESIDENT** Barbara A. Schaal **PRESIDENT-ELECT** Susan Hockfield **TREASURER** David Evans Shaw **CHIEF EXECUTIVE OFFICER** Rush D. Holt **BOARD** Cynthia M. Beall, May R. Berenbaum, Carlos J. Bustamante, Stephen P.A. Fodor, Claire M. Fraser, Michael S. Gazzaniga, Laura H. Greene, Elizabeth Loftus, Mercedes Pascual

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417 FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES 202-326-6730 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Services** 202-326-6417 or <http://membercentral.aaas.org/discounts>

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority of conflicting points of view. Accordingly, all articles published in Science—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

INFORMATION FOR AUTHORS See pages 624 and 625 of the 5 February 2016 issue or access www.sciencemag.org/authors/science-information-authors

SENIOR EDITORIAL BOARD

Gary King, *Harvard University*, Susan M. Rosenberg, *Baylor College of Medicine*, Ali Shilatfard, *Northwestern University Feinberg School of Medicine*

BOARD OF REVIEWING EDITORS

(Statistics board members indicated with \$)

Adriano Aguzzi, *U. of Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Judith Allen, *U. of Edinburgh*
Sonia Altizer, *U. of Georgia*
Sebastian Amigorena, *Institut Curie*
Kathryn Anderson, *Memorial Sloan-Kettering Cancer Center*
Meinrat O. Andreae, *Max-Planck Inst. Mainz*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Clare Baker, *University of Cambridge*
Nenad Ban, *ETH Zürich*
Jordi Bascompte, *University of Zurich*
Franz Bauer, *Instituto de Astrofísica*
Ray H. Baughman, *U. of Texas, Dallas*
David Baum, *U. of Wisconsin*
Carlo Beenakker, *Leiden U.*
Kamran Behnia, *ESPCI-ParisTech*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
May Berenbaum, *U. of Illinois*
Gabriele Bergers, *U. of California, San Francisco*
Bradley Bernstein, *Massachusetts General Hospital*
Peer Bork, *EMBL*
Bernard Bourdon, *Ecole Normale Supérieure de Lyon*
Chris Bowler, *Ecole Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *U. of California, Santa Cruz*
Ron Brookmeyer, *U. of California Los Angeles (\$)*
Christian Büchel, *U. Hamburg-Eppendorf*
Joseph A. Burns, *Cornell U.*
Carter Tribley Butts, *U. of California, Irvine*
Gyorgy Buzsaki, *New York U. School of Medicine*
Blanche Capel, *Duke U.*
Mats Carlsson, *U. of Oslo*
Ib Chorkendorff, *U. of Denmark*
David Clapham, *Children's Hospital Boston*
Joel Cohen, *Rockefeller U., Columbia U.*
James J. Collins, *MIT*
Robert Cook-Deegan, *Duke U.*
Lisa Coussens, *Oregon Health & Science U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Robert H. Crabtree, *Yale U.*
Roberta Croce, *Vrije Universiteit*
Janet Currie, *Princeton U.*
Jeff L. Dangel, *U. of North Carolina*
Tom Daniel, *U. of Washington*
Frans de Waal, *Emory U.*
Stanislas Dehaene, *Collège de France*
Robert Desimone, *MIT*
Claude Desplan, *New York U.*
Dennis Discher, *U. of Pennsylvania*
Gerald W. Dorn II, *Washington U. School of Medicine*
Jennifer A. Doudna, *U. of California, Berkeley*
Bruce Dunn, *U. of California, Los Angeles*
William Dunphy, *Caltech*
Christopher Dye, *WHO*
Todd Ehlers, *U. of Tuebingen*
David Ehrhardt, *Carnegie Inst. of Washington*
Tim Elston, *U. of North Carolina at Chapel Hill*
Jennifer Elisseeff, *U. of N.C.*
Gerhard Ertl, *Fritz-Haber-Institut, Berlin*
Barry Everitt, *U. of Cambridge*
Ernst Fehr, *Johns Hopkins U.*
Anne C. Ferguson-Smith, *U. of Cambridge*
Michael Feuer, *The George Washington U.*
Toren Finkel, *NHLBI, NIH*
Kate Fitzgerald, *U. of Massachusetts*
Peter Fratzl, *Max-Planck Inst.*
Elaine Fuchs, *Rockefeller U.*
Daniel Geschwind, *UCLA*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *Rice U.*
Julia R. Greer, *Caltech*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zurich*
Kip Guy, *St. Jude's Children's Research Hospital*
Taekjip Ha, *U. of Illinois at Urbana-Champaign*
Wolf-Dietrich Hardt, *ETH Zurich*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *U. of Illinois at Urbana-Champaign*
Michael Hasselmo, *Boston U.*
Martin Heimann, *Max-Planck Inst. Jena*
Yka Helariutta, *U. of Cambridge*
James A. Hendler, *Rensselaer Polytechnic Inst.*
Janet G. Hering, *Swiss Fed. Inst. of Aquatic Science & Technology*
Kai-Uwe Hinrichs, *U. of Bremen*
David Hodell, *U. of Cambridge*
Lora Hooper, *UT Southwestern Medical Ctr. at Dallas*
Tamas Horvath, *Yale University*
Raymond Huey, *U. of Washington*
Fred Hughson, *Princeton U.*
Auke Ijspeert, *EPFL Lausanne*
Stephen Jackson, *USGS and U. of Arizona*
Steven Jacobsen, *U. of California, Los Angeles*
Kai Johnson, *EPFL Lausanne*
Peter Jonas, *Inst. of Science & Technology (IST) Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kahne, *Harvard U.*
Daniel Kammen, *U. of California, Berkeley*
Abby Kavner, *U. of California, Los Angeles*
Hitoshi Kawakatsu, *U. of Tokyo*
Masashi Kawasaki, *U. of Tokyo*
V. Narry Kim, *Seoul National U.*
Robert King, *Harvard Medical School*
Etienne Koechlin, *Ecole Normale Supérieure*
Alexander Kolodkin, *Johns Hopkins U.*
Thomas Langer, *U. of Cologne*
Mitchell A. Lazar, *U. of Pennsylvania*
David Lazer, *Harvard U.*
Thomas Lecuit, *IBDM*
Virginia Lee, *U. of Pennsylvania*
Stanley Lemon, *U. of North Carolina at Chapel Hill*
Ottoline Leyser, *Cambridge U.*
Wendell Lim, *U.C. San Francisco*
Marcia C. Linn, *U. of California, Berkeley*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzan, *CIC biomaGUNE*
Jonathan Losos, *Harvard U.*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Laura Machesky, *CRUK Beatson Inst. for Cancer Research*
Anne Magurran, *U. of St. Andrews*
Oscar Marin, *CSIC & U. Miguel Hernández*
Charles Marshall, *U. of California, Berkeley*
C. Robertson McClung, *Dartmouth College*
Rodrigo Medellín, *U. of Mexico*
Graham Medley, *U. of Warwick*
Tom Misteli, *NCI*
Yasushi Miyashita, *U. of Tokyo*
Mary Ann Moran, *U. of Georgia*
Richard Morris, *U. of Edinburgh*
Alison Mutsaers-Reif, *NC State U. (\$)*
Thomas Murray, *The Hastings Center*
Daniel Neumark, *U. of California, Berkeley*
Kitty Nijmeijer, *U. of Twente*
Helga Nowotny, *European Research Advisory Board*
Ben Olken, *MIT*
Rachel O'Reilly, *Warwick U.*
Joe Orenstein, *U. of California Berkeley & Lawrence Berkeley National Lab*
Harry Orr, *U. of Minnesota*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Margaret Palmer, *U. of Maryland*
Steve Palumbi, *Stanford U.*
Jane Parker, *Max-Planck Inst. of Plant Breeding Research*
Giovanni Parmigiani, *Dana-Farber Cancer Inst. (\$)*
John H. J. Petrini, *Memorial Sloan-Kettering Cancer Center*
Samuel Pfaff, *Salk Institute for Biological Studies*
Joshua Plotkin, *U. of Pennsylvania*
Albert Polman, *FOM Institute AMOLF*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Wim van der Putten, *Netherlands Institute of Ecology*
David Randall, *Colorado State U.*
Felix Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Jim Roberts, *Fred Hutchinson Cancer Research Ctr.*
Amy Rosenzweig, *Northwestern University*
Mike Ryan, *U. of Texas, Austin*
Mitinori Saitou, *Kyoto U.*
Shimon Sakaguchi, *Kyoto U.*
Miguel Salmeron, *Lawrence Berkeley National Lab*
Jürgen Sandkühler, *Medical U. of Vienna*
Alexander Schier, *Harvard U.*
Vladimir Shalae, *Purdue U.*
Robert Siliciano, *Johns Hopkins School of Medicine*
Denis Simon, *Arizona State U.*
Uri Simonsohn, *U. of Pennsylvania*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina (\$)*
John Speakman, *U. of Aberdeen*
Allan C. Spradling, *Carnegie Institution of Washington*
Jonathan Sprent, *Garvan Inst. of Medical Research*
Eric Steig, *U. of Washington*
Paula Stephan, *Georgia State U. and National Bureau of Economic Research*
Molly Stevens, *Imperial College London*
V. S. Subrahmanian, *U. of Maryland*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Cambridge U.*
John Thomas, *North Carolina State U.*
Shubha Tole, *State Institute of Fundamental Research*
Christopher Tyler-Smith, *The Wellcome Trust Sanger Inst.*
Herbert Virgin, *Washington U.*
Bert Vogelstein, *Johns Hopkins U.*
Cynthia Volkert, *U. of Göttingen*
David Wallace, *Weizmann Inst. of Science*
Ian Walsmsley, *U. of Oxford*
Jane-Ling Wang, *U. of California, Davis (\$)*
David Waxman, *Fudan U.*
Jonathan Weissman, *U. of California, San Francisco*
Chris Willmott, *U. of Missouri (\$)*
Ian A. Wilson, *The Scripps Res. Inst. (\$)*
Timothy D. Wilson, *U. of Virginia*
Rosemary Wyse, *Johns Hopkins U.*
Jan Zaenen, *Leiden U.*
Kenneth Zaret, *U. of Pennsylvania School of Medicine*
Jonathan Zehr, *U. of California, Santa Cruz*
Len Zon, *Children's Hospital Boston*
Maria Zuber, *MIT*

BOOK REVIEW BOARD

David Bloom, *Harvard U.*, Samuel Bowring, *MIT*, Angela Creager, *Princeton U.*, Richard Swedder, *U. of Chicago*, Ed Wasserman, *DuPont*

EIT's teething problems

Good policy intentions do not always translate into the desired outcome. Nearly a decade ago, the European Commission (EC) launched two new entities with the goal of building a European-wide ecosystem—one to support frontier research, the European Research Council (ERC), and the other to boost innovation, the European Institute of Technology (EIT). Today, the ERC is considered the gold standard for funding frontier research of scientific excellence, with a budget that almost doubled in 2014 under Horizon 2020, the European Union Framework Programme for Research and Innovation. What about the EIT? While acknowledging the progress made, an April report by the European Court of Auditors states that the EIT's overly complex operational structure and past management problems have impeded its overall effectiveness. In fact, continuous evaluations and adjustments will probably be needed to turn it into a success.

In 2005, EC President José Manuel Barroso proposed the EIT as a European beacon for facilitating innovation. The EIT should become a rallying point for industry, research, and education across the extremely diverse innovation landscape in Europe. Reservations arose immediately about the EIT's fit into already existing structures and schemes, inadequate funding, and the time required for such an entity to mature. The latter proved particularly true.

Although both the ERC and EIT had to navigate the bureaucratic maze of EC rules, the EIT has experienced considerably more growing pains. This is not surprising, given that it aims to create an innovation ecosystem, which is a far more arduous task than supporting frontier research. As it turned out, the EIT's structure would mirror the complexities to be confronted. Its main operative instruments are competitively selected Knowledge and Innovation Communities (KICs), which nurture entrepreneurial students through EIT-labeled study programs at universities. In addition to this academic arm, a wealth of EIT activities foster joint research between business and

academia, create innovative companies, and support their growth. To accomplish all of these missions, KICs were set up as legal entities, giving them flexibility to tailor support for the needs of entrepreneurs and companies. But KICs were also intended to develop strategies for financial sustainability. Given that the European innovation system is dominated by a multitude of granting schemes for different purposes, this goal is ambitious, if not unrealistic.



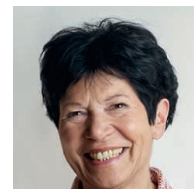
“...EIT holds important lessons for the future.”

In its report, the auditors in particular expressed doubts about the financial sustainability of the KICs. They also found the mechanism of annual grant cycles ill suited for the KICs' multiannual activities. They urged several regulatory and operational adjustments of the EIT if expectations are to be fulfilled. For example, EIT funding may only cover a maximum of 25% of a KIC's overall costs, whereas complementary activities are covered from other sources, even ERC grants. The auditors found that there is little added value to this funding condition, which unnecessarily complicates reporting.

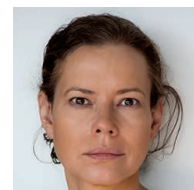
Arguably, igniting a new entrepreneurial spirit across universities stands out as the EIT's most substantial contribution thus far. Nevertheless, the search continues for an instrument that can address innovation as successfully as the ERC supports excellent science. In 2014, the EC launched the Small and Medium-sized Enterprises instrument to fund high-potential innovation. Earlier this year, a public consultation for setting up a European Innovation Council, inspired by the success of the ERC, was carried out by Research, Science and Innovation Commissioner Carlos Moedas.

It is too early to know the outcome of these attempts to boost Europe's performance in innovation, or how they will complement the EIT. In the meantime, continuous monitoring, analysis, and adaptation of the EIT's efforts may lead to a uniquely European facilitator of innovation. It is clear that no single, magic solution exists. As a part of the larger innovation ecosystem still under construction, the EIT holds important lessons for the future.

—Helga Nowotny and Jana Kolar



Helga Nowotny is former president of the ERC and professor emerita at ETH, Zurich, Switzerland. Email: helga.nowotny@wtf.at



Jana Kolar is executive director of Central European Research Infrastructure Consortium and a member of the governing board of the EIT. Email: jana.kolar@ceric-eric.eu

“I believe the letter was an attempt at intimidation; my deans certainly interpreted it as such.”

Gabriel Gardner, a librarian at California State University, Long Beach, to Inside Higher Ed, after the Association of American Publishers complained to the university that Gardner's research on pirated paper repositories such as Sci-Hub served as an endorsement.

IN BRIEF

Scientists celebrate bird treaty



A tracking tag antenna pokes through the feathers of a gray catbird.

A field station popped up last week amid the famous monuments framing Washington, D.C.'s National Mall. Ornithologists set up nets to briefly capture some of the park's flying residents, including gray catbirds, song sparrows, and mourning doves. But the scientists' real objective was to draw attention to a landmark in bird conservation, the 100th anniversary of the Migratory Bird Treaty Act. The act, an agreement between the United States and Canada, stemmed in part from ornithologists' warnings that unregulated hunting was decimating many North American bird populations, says Pete Marra, head of the Smithsonian Migratory Bird Center in Washington, D.C. Mexico later joined the pact, and the United States also struck deals with Russia and Japan. Such international collaboration is needed now more than ever, with dozens of species of migratory birds facing dramatic population declines, Marra said on 3 August at the event on the mall. Some migratory birds alight in six or more nations during a year, he noted, so no one country can protect them. To aid in the effort, scientists are working to better map bird migrations and understand habitat use with the help of technologies such as microelectronics to create tiny, lightweight tracking devices. http://scim.ag/_MBA100

AROUND THE WORLD

Company OK'd to visit moon

CAPE CANAVERAL, FLORIDA | A company called Moon Express, based in Cape Canaveral, Florida, on 3 August became the first private company to receive permission from the U.S. Federal Aviation Administration (FAA) to travel out of Earth's orbit. Moon Express, formed in 2010 to enter the \$30 million Google Lunar XPRIZE competition, now has a green light to send a robotic spacecraft to the surface of the "eighth continent" in 2017. In October 2015, the group had signed a contract with aerospace company Rocket Lab to send three missions to the moon between 2017 and 2020. In April, Moon Express submitted its request to FAA, which ultimately determined that its payload launch "does not jeopardize public health and safety, safety of property, U.S. national security or foreign policy interests, or international obligations of the United States." The company last month also signed an agreement to assemble and test its lander at Cape Canaveral.

Moving an endangered tortoise

PERTH, AUSTRALIA | As long as it has been known to science, the diminutive western swamp tortoise has been in peril. Presumed extinct in 1901, it was rediscovered in the 1950s; since then, biologists have struggled to protect it from habitat loss and introduced predators. Now, with declining rainfall due to climate change posing an even more urgent threat, scientists have devised a controversial plan to



The western swamp tortoise will get a new home.

safeguard the animal's future: On 11 August, they released the tortoises to new sites outside of their known historical range, making them the first vertebrates to be deliberately relocated because of climate change. The relocation begins a yearlong trial that will track 24 juvenile tortoises released to two sites roughly 250 kilometers south of the animals' native habitat near Perth, Australia. Although the sites aren't ideal for the tortoises now, detailed modeling of rainfall, temperature, swamp hydrology, and tortoise biology predict they will be in half a century. But some scientists worry that the impacts of such "planned invasions" are difficult to predict, and may set an alarming precedent for other species. <http://scim.ag/WSTrelocate>

About-face on climate research

CANBERRA | Australia's new science minister has ordered the nation's premier science agency to "put the focus back on climate science"—and Australian scientists have their fingers crossed, hoping the directive from Greg Hunt, revealed 4 August, really indicates the federal government is reversing a previous decision to scale back climate research efforts. In February, the government announced a realignment of priorities for the Commonwealth Scientific and Industrial Research Organisation (CSIRO) that called for eliminating 350 jobs, including 110 climate science positions (*Science*, 12 February, p. 649). The agency later scaled back the job cuts to 295 positions, including more than 60 climate and marine scientists. In the recent announcement, Hunt said he would work with Australia's chief scientist, Alan Finkel, CSIRO, and the scientific community to develop a new climate science strategy that will cover remaining staff and charge a new climate change center in Hobart, Australia, with coordinating related research efforts across government agencies and academia. The Australian Academy of Science welcomed Hunt's statement, as did the union representing CSIRO staff. But CSIRO Staff Association Secretary Sam Popovski in Melbourne said the "backflip" on climate science capacity was "merely a Band-Aid" given that the agency is proceeding with the staff cuts. <http://scim.ag/Austclimate>

Stretching yellow fever vaccines

KINSHASA | An emergency measure designed to stretch supplies of yellow fever vaccine is about to be used for the first time. In a mass vaccination campaign



Hoaxer Charles Dawson (second from right, back row), and others study the Piltdown Man's skull in this 1915 painting.

A lone hoaxer dreamed up Piltdown Man

In 1912, amateur fossil hunter Charles Dawson stunned scientists at a meeting of the Geological Society of London when he claimed to have found fossils belonging to a significant "missing link" in human evolution. The fossils—some skull fragments, teeth, and a jawbone, along with several stone tools—were found in a gravel pit near the U.K. village of Piltdown. The Piltdown Man, said to be about 500,000 years old, had an unusually thick, humanlike cranium, an apelike mandible, and canine teeth somewhere between a human's and an ape's in size. It shocked the evolutionary research community at the time—and, scientists discovered some 40 years later, it was a total fabrication. The Piltdown Man was actually cobbled together from two human skulls and an orangutan's jaw and teeth. Yet questions remained over which of the various people who dug up and investigated the bones were involved in the deceit. (Even Sherlock Holmes creator Arthur Conan Doyle was considered at one point.) Now, a new study published in *Royal Society Open Science* suggests that a single hoaxer, likely Dawson, acted alone. The Piltdown Man's bones were the handiwork of one person who laboriously filed down teeth, stained bones, and filled them with loose gravel to make them heavier, the researchers conclude. "Throughout the whole assemblage, there's evidence of one hand, one maker, one signature," says paleoanthropologist Isabelle De Groote of Liverpool John Moores University in the United Kingdom, who led the study.

scheduled to start next week in Kinshasa, the capital of the Democratic Republic of the Congo (DRC), more than 8 million people are going to be vaccinated using just one-fifth the normal dose. Yellow fever vaccine is in short supply and the outbreak, which started in Angola in December 2015 and has since spread to the DRC, has depleted global stocks. The DRC has so far reported 2051 suspected cases including 95 deaths from the mosquito-borne disease. The Strategic Advisory Group of Experts on Immunization at the World Health Organization had concluded in June that the smaller dose would still offer protection for at least 12 months (*Science*,

24 June, p. 1498). It is unclear, however, how much longer that protection might last, and scientists are pushing for research to be done on the effects of the emergency dose. <http://scim.ag/yellowfevvac>

FINDINGS

How horses' gaits were smoothed

Horses' strides became a lot smoother about 12 centuries ago. A single genetic mutation, which arose in the medieval United Kingdom and Iceland, gave horses their ability to "amble," or walk with a relatively smooth, four-beat rhythm versus



A U.N. group proposed making the Sargasso Sea a "high seas" World Heritage Site.

World Heritage Sites in the high seas

A U.N. group is urging the world's governments to expand a longstanding—but largely symbolic—conservation strategy to international waters. To protect important high seas ecosystems, governments should establish ways to designate World Heritage Sites in the two-thirds of the world's seas outside national borders, argues a report released last week by the United Nations Educational, Scientific and Cultural Organization (UNESCO). The 1972 World Heritage Convention, ratified by nearly 200 nations, allows the nations to nominate places of special cultural or physical significance to a list of UNESCO World Heritage Sites. The list includes more than 1000 sites, from India's Taj Mahal to Tanzania's Serengeti National Park, as well as 47 marine areas such as Ecuador's Galápagos Islands and Australia's Great Barrier Reef. Protecting the sites is largely left up to national governments; the 1972 agreement, however, didn't establish a clear way to designate sites that are not within national boundaries. The time has come to resolve that problem, the report argues; it identifies five high seas sites for World Heritage designation, including a Pacific Ocean "thermal dome" near Costa Rica where many endangered species feed and spawn, the Atlantic's famed Sargasso Sea, and the Indian Ocean's Atlantis Bank—a sunken island full of rare fossils. <http://scim.ag/HeritageSeas>

a bumpier, more erratic pattern, researchers reported this week in *Current Biology*. Scientists first detected a mutation in the gene *DMRT3*, which influences limb movements, in all ambling horses in 2012. But they weren't sure how the mutation had changed the horses' gait. Studies in mice now suggest that the *DMRT3* mutation affects the development of the spinal cord and allows longer strides. To trace its spread in horses, researchers extracted DNA from the bones of 90 ancient horses that lived as far back as 3500 B.C.E., tracing the rise of the mutation to roughly 850 to 900 C.E. The scientists speculate that Vikings brought horses from the United Kingdom to Iceland and bred them together around this time; they then disseminated those horses to mainland Europe. Riders' preference for the comfortable rides of ambling horses helped the mutation quickly spread to horses worldwide, they suggest.

NEWSMAKERS

Three Qs

On 1 October, **Ian Chapman** takes the helm of the Culham Centre for Fusion Energy (CCFE) in Abingdon, the United Kingdom's national fusion lab. Chapman, a physicist, has worked at the lab since joining as a grad student in 2004. He spoke with *Science* about CCFE's innovative fusion reactor, the Mega Amp Spherical Tokamak (MAST), the Joint European Torus (JET)—the world's biggest fusion reactor until the International Thermonuclear Experimental Reactor (ITER) project is completed in 2025—and how the lab will continue to collaborate with Europe after Brexit.

Q: MAST was recently upgraded. What does CCFE hope to achieve with it?

A: MAST-Upgrade is the most innovative and exciting machine in fusion at the moment. It will give us an idea of how to

exhaust heat from a fusion reactor in a more efficient way, leading to smaller, cheaper reactors.

Q: Will JET continue its work in support of ITER?

A: JET is definitely the best place to prepare for ITER. We can prepare the physics basis [for ITER operation] so it can achieve its highest performance as quickly as possible. We're in negotiations with ITER partners to extend JET into the 2020s. Everyone wants to continue and we'll have to thrash out how it works, but Brexit doesn't simplify things.

Q: What is the future of U.K. fusion post-Brexit?

A: We must find a way for the United Kingdom to remain at the forefront and collaborate. You learn from collaborations and the experiences you share, and that makes you stronger together.

BY THE NUMBERS

12%

Reduction in honey bee colonies in 29 countries after the winter of 2015–16, with the heaviest losses in the United Kingdom and Spain, according to the nonprofit honey bee research network Prevention of Honey Bee Colony Losses.

1700

Year of the first recorded usage of a microscope to perform an autopsy—on the body of Pope Innocent XII, who died that year (*The Lancet*).

26%

Percent fewer violent infractions committed by inmates at a prison who watched nature videos several times a week over a year, relative to inmates who didn't watch the videos, scientists reported last week at the American Psychological Association's annual meeting.



EDUCATION

Founder's death unsettles Egypt's science city

Nobel Prize-winning chemist Ahmed Zewail was key to fundraising and recruitment

By Robert F. Service

Nobel Prize-winning chemist Ahmed Zewail lived to see his dream take shape: an elite research university, housed within a science city, in his native Egypt. The fledgling organization, Zewail City of Science and Technology on the outskirts of Cairo, began admitting students to its university in 2013. "It's really a beautiful story of hope and a new beginning for Egypt," Zewail said this past April in an interview with *Science*. Last week, however, Zewail's death at the age of 70 deprived the effort of its fundraiser-in-chief and most powerful advocate.

Leaders of the project have vowed to press on with their plans to build a center for science and innovation in the Arab world. But some have concerns about keeping tuition costs low and completing fundraising goals. "There is a worry that if his name is not there, [fundraising] could be affected," says Mostafa El-Sayed, a chemist at the Georgia Institute of Technology in Atlanta and a member of Zewail City's advisory council. "Let's hope not."

Zewail, who was born in Egypt's Nile delta but spent most of his career at the California Institute of Technology in Pasadena, first proposed the

university in 1999. That was the year he won the chemistry Nobel Prize for pioneering the field of femtochemistry: using laser pulses as a stop action camera to track chemical reactions as they unfold. As the first Arab to win a Nobel Prize in the sciences, the naturalized U.S. citizen gained enormous prestige—serving, among other things, as the first U.S. science envoy to the Middle East aimed at improving U.S. relations with the Muslim world.

It had always bothered Zewail that the best and brightest Egyptian students typically had to leave the country to advance their careers. He hoped that the creation of a top-flight research university would turn things around and lure expatriate

Egyptian scientists back home. In 1999, he sketched out his ideas to former Egyptian President Hosni Mubarak, who initially embraced the plan.

A key element of Zewail's vision was difficult for Egypt's autocratic leader to swallow, however. At most of the country's research institutions, the federal government holds sway over university faculty hiring decisions and research funds. But Zewail wanted to promote an independent, merit-based scientific culture in Egypt. He insisted that the university be run by an independent board of trustees with complete authority over hiring and admission decisions. "This was very irritating to Mubarak," Zewail recalled. Without the president's backing, the project stalled. But shortly after Mubarak was ousted in the Egyptian revolution of 2011, interest in the project revived. Six months later, the Egyptian cabinet declared Zewail City for Science and Technology a national project.

The city is designed to include teaching facilities for about 1000 students, seven thematic research institutes modeled on the Max Planck Institutes in Germany, and a technology transfer center aimed at propelling research advances into startup companies. In 2014, the Egyptian government gave the city 80 hectares for the campus and, a year later, said it would



Ahmed Zewail, 70, died last week, having seen his namesake science city nearly to fruition.

Zewail City of Science and Technology's university began admitting students, tuition-free, in 2013.

pay for the construction costs. The university's first class is due to graduate next year, and, for the initial round of students, the tuition has been free.

Zewail managed to recruit Egyptian scientists abroad to return to head most of the research institutes, which focus on research themes such as nanotechnology, energy, and biomedicine. Next year, the city is set to complete phase one of its building projects with 10 teaching and research facilities. According to Ashraf Badawi, Zewail City's dean of student affairs, the science city is still planning and fundraising for a second phase, which calls for an additional four research buildings. Eventually, the university could grow to as many as 5000 students, Zewail said in April. But specific plans for such a large expansion are yet to be drafted.

Prior to his death, Zewail said that he had already helped raise nearly \$1 billion of the project's eventual \$2 billion goal. Much of that came through Zewail's personal contacts with business and political leaders both inside Egypt and abroad, as well as direct appeals to the Egyptian public. "He knew everyone," Badawi says, and encouraged them all to contribute to the project. In 2011, for example, an Egyptian pharmaceutical executive, Hassan Abbas Helmy, gave 250 million Egyptian pounds (\$42 million) for a pharmaceutical research institute.

"We are hoping this will continue" despite Zewail's death, Badawi says. In any case, he says, there's enough funding in place for the foreseeable future. "The city should be fine for several years." At Zewail's funeral, held with full military honors on 7 August in a Cairo suburb, Egyptian President Abdel-Fattah el-Sisi insisted the government would see the project through to completion "one way or another," according to an article in *The New York Times*.

Others at Zewail City's university say they are shaken by Zewail's passing but determined to sustain his vision. "He meant a lot to us," says Reem Khidr Arafa, a U.S.-trained professor of biomedical sciences at Zewail City whom Zewail recruited from Cairo University to join the faculty in 2014. Shortly after learning the news, Arafa says, the university faculty had a brief meeting. "Everyone was in shock," Arafa says. "Now, everyone feels the responsibility to continue the mission. We believe in his dream." That dream, Zewail said back in April, "is a new experiment for Egypt. We need a couple of more years to finish the job. But we are moving in the right direction." ■

CONSERVATION BIOLOGY

Can captive breeding save Mexico's vaquita?

Scientists mull a risky strategy to save an imperiled porpoise

By Ben Goldfarb

Species don't come much more endangered than the vaquita, a child-sized porpoise that is threatened by fishing nets in the northern reaches of Mexico's Gulf of California. Just 60 remain, experts warned earlier this year, solidifying *Phocoena sinus*'s status as the world's most endangered marine mammal. That grim assessment now has researchers pondering a controversial strategy: capturing a handful of vaquitas and breeding them in captivity.

"Given the crisis we're in, we need to explore all of our options," says biologist

risk of killing a vaquita while catching them is very high. With only 50 or 60 animals left, we can't play with that."

The population of vaquitas, the world's smallest cetaceans at 1.5 meters long, has been declining by an estimated 34% annually since 2011, almost entirely because of fishing with gillnets, which entangle and drown the animals. In April 2015, the Mexican government imposed a temporary 2-year ban on gillnets within the vaquita's range, and on 22 July it made the ban permanent, a move long recommended by CIRVA's scientists. Illegal nets still pose a threat, however, as poachers pursue a fish called the totoaba, whose bladder fetches up to \$20,000 in



Set to capture sharks and other fish, a gillnet also snared a vaquita, a small, endangered porpoise.

Barbara Taylor of the National Oceanic and Atmospheric Administration's Southwest Fisheries Science Center in San Diego, California, who serves on the International Committee for the Recovery of the Vaquita (CIRVA). "Keeping some individuals in a sanctuary is one of those options."

The idea is fraught with practical and political difficulties. No one has ever tried to capture, transport, or care for the animals. And some conservationists fear a captive breeding program will undermine efforts to save the species in the wild. "I don't like this idea at all," says Omar Vidal, director general of the environmental group World Wildlife Fund (WWF) Mexico in Mexico City. "The

China. Poachers killed at least three vaquitas this past March alone.

Given the continuing danger, researchers say they must consider captive breeding. The strategy has helped save the black-footed ferret of the western prairies and the California condor, whose populations had both dwindled to about 20 animals. But it has never been tried with a cetacean, and although some small marine mammals, such as bottlenose dolphins, thrive in captivity, porpoises are ill-suited for confinement. "They're very sensitive to stress and noise, and they have high heart rates," says Frances Gulland, a CIRVA member and senior scientist at the Marine Mammal Center in Sausalito,

California. “We think of them as the hummingbirds of the marine mammal world.”

In recent years, however, researchers in Denmark, the Netherlands, and elsewhere have shown that some porpoises can be captured and kept in captivity. Those advances convinced Gulland that it was worth exploring the strategy with the vaquita, and last year she convened an expert team in the Netherlands to assess the idea’s feasibility.

Each phase of the process would pose challenges, the team concluded. Locating the elusive cetaceans in the choppy, murky waters of the gulf is tricky; the team even suggested exploring the use of trained bottlenose dolphins, such as those kept by the U.S. Navy, to help with the hunt. Once located, the team envisions herding vaquitas into lightweight surface gillnets, which are safer than conventional nets—a tactic that’s been used to place satellite tags on harbor porpoises in Greenland. Then, they would use a moist stretcher to transport each animal to a soft-sided net pen, and likely later a large artificial pool, along the gulf’s coast. There, they would figure out how to feed and care for the animals, and attempt to persuade them to breed. The ultimate goal would be to release some parents and offspring back into the wild once the threat of gillnets has been mitigated.

The first vaquita capture could occur in 2017, if further study supports the idea, a recent CIRVA report notes. The likely first target would be a young male, Gulland says, because a loss in that demographic group would be least harmful to the population if things go awry.

WWF’s Vidal fears, however, that the strategy could take the pressure off Mexican authorities to crack down on illegal fishing. “Species need to recover in the wild,” he adds. Vidal notes that Mexico’s Guadalupe fur seal, which was hunted nearly to extinction in the 19th century, has rebounded without captive breeding.

CIRVA biologists acknowledge the strategy’s limitations. “There’s no point in putting vaquitas into a sanctuary if they’re just going to be killed once you release them,” Gulland says. But they say capturing a few vaquitas—rather than the entire population, as was done with condors and ferrets—could provide an insurance policy against extinction.

If history is any guide, they’ll have to act fast. In 2006, Taylor was on a team that hoped to capture baiji, an endangered freshwater dolphin that lived in China’s Yangtze River, and relocate the animals to protected lakes. But it was too late: The researchers never found any baiji, and the species was declared functionally extinct. ■

Ben Goldfarb is a freelance writer in New Haven, Connecticut.

BIOETHICS

NIH plans to fund human-animal chimera research

Agency would subject experiments—such as growing human organs in pigs—to extra ethical review

By Jocelyn Kaiser

Last September, in a move that took researchers by surprise, the National Institutes of Health (NIH) announced it would not fund controversial experiments that add human stem cells to animal embryos. Now, the agency says it is ready to move forward with some so-called chimera studies as long as they pass a special agency ethics review. In a draft policy released last week, NIH said it will take an especially close look at studies that add human cells to very early embryos of certain animals or insert them into an animal’s brain.

Biomedical scientists generally welcomed the proposal, which could open the way to growing human organs in animals or studying human brain disease in monkeys. “I applaud the actions NIH is taking to advance this area of research in a responsible and timely manner,” says developmental biologist Juan Carlos Izpisua Belmonte of the Salk Institute for Biological Studies in San Diego, California, whose application for a prestigious NIH research award was put on hold last fall because of the moratorium on chimera studies. But some were left trying to parse exactly what NIH’s policy will mean. “We still don’t know what the out-

come will be case by case,” says Sean Wu, a stem cell researcher at Stanford University in Palo Alto, California, who co-authored a letter to *Science* last year opposing the moratorium.

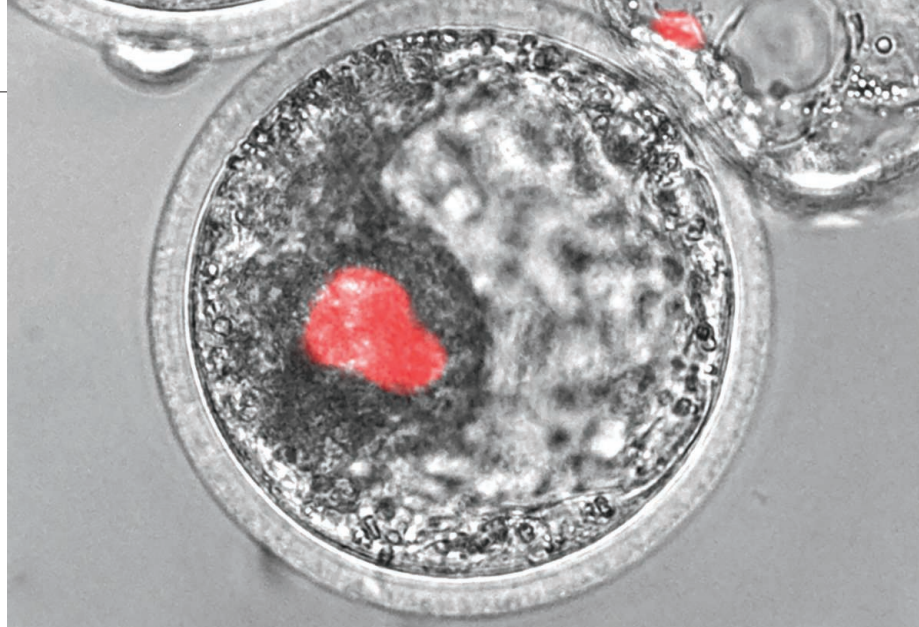
At issue are experiments in which scientists introduce human pluripotent stem cells—cells that can turn into any kind of tissue—into early embryos of mice and other laboratory animals and then let the animals develop. Such experiments can be used to study human development, generate disease models, and potentially grow human organs and tissues for transplantation. Several scientists, including Izpisua Belmonte, have already begun using non-NIH research funds to inject stem cells into pig or sheep embryos, in an effort to grow human pancreases or other organs inside the animals. But the public has been leery of such experiments, with some scientists and ethicists worrying that they could produce an animal with humanlike cognitive abilities or lead to a human embryo growing inside an animal. Several countries already restrict this research by law (see table, below).

Responding to growing interest in using chimeras to produce human organs, NIH decided last fall to suspend reviews of new funding applications while the agency

The legal landscape for chimera studies

Most countries don’t explicitly permit or forbid human-animal chimera research, but a few have relevant laws or policies on funding for such work.

COUNTRY	POLICY
France	Law forbids creating a chimeric human embryo, but is less clear on whether adding human cells to animal embryos is allowed.
United Kingdom	Regulations issued in January require extra ethical review for human-animal chimera experiments that involve nonhuman primate cells, germ cells, or the brain, or that affect an animal’s appearance or behavior.
Germany	Law forbids combining a human embryo with animal cells, but not the introduction of human cells into an animal embryo.
Japan	Law limits research on human-animal chimeric embryos, not allowing development beyond the appearance of the primitive streak or transfer into an animal. A bioethics panel recently proposed more permissive, case-by-case review.
United States	No legal prohibition, but advisory bodies have recommended limits on breeding chimeras and adding human cells to primate embryos. After a funding pause, the National Institutes of Health now proposes case-by-case review.



The National Institutes of Health may now fund experiments such as this one in which human cells (red) are added to a pig embryo.

considered the ethical issues (*Science*, 16 October 2015, p. 261). (It was not funding any such chimera studies at the time.) NIH then held a workshop last November to gather input from the research community and bioethicists, where most argued that these studies are scientifically valuable.

According to its 4 August announcement, NIH proposes to lift the moratorium but impose a new agency review process on certain especially sensitive experiments. One type involves adding human stem cells to very early vertebrate embryos—through the gastrulation stage, when an embryo develops three distinct layers of cells that then give rise to different tissues and organs. The other category needing extra review would be studies that introduce human cells into the brains of postgastrulation mammals. (The latter kind of experiment would not need the extra review if done in rodents, which are considered unlikely to acquire humanlike capabilities.)

Along with undergoing traditional scientific peer review, proposed studies of either type would go to an internal NIH steering committee of scientists, ethicists, and animal welfare experts. They would consider factors such as the type of human cells, where they may wind up in the animal, and how the cells might change the animal's behavior or appearance. The committee's conclusions would then help NIH's institutes decide whether to fund the projects.

NIH also wants to tighten its existing stem cell guidelines to prohibit studies that add human stem cells to primate embryos at any point up to and including the pregastrulation stage known as the blastocyst. (Current guidelines only prohibit adding human pluripotent cells to primate blastocysts.) And

the agency wants to extend a current ban on breeding chimeric animals that might carry human eggs or sperm. Using private funds to create chimeras of any kind would remain legal in the United States, unlike in several countries.

"I am confident that these proposed changes will enable the NIH research community to move this promising area of science forward in a responsible manner," wrote Carrie Wolinetz, associate director for science policy at NIH in Bethesda, Maryland, in a blog post. On a call with reporters, she emphasized that the proposal "is not ... a prohibition" on chimera research. "It is merely an extra look." NIH is collecting comment on the proposed changes until 4 September; it hopes to issue a final policy and lift the moratorium by late January 2017, Wolinetz said.

The proposal could open up research in areas that had been potentially off-limits. Neuroscientist Steve Goldman of the University of Rochester in New York notes that even experiments that put human cells into the brains of monkeys or other primates after the blastocyst stage could be allowed. Such studies, which could be useful for studying mental illnesses, "had been a very murky zone" until now, Goldman says. The proposed changes suggest "a much more permissive environment."

But some researchers have qualms. Cell biologist Stuart Newman of New York Medical College in Valhalla, a longtime opponent of chimera research who opposes lifting the moratorium, says he realizes he is among few dissenters in the research community. It's not the idea of growing human organs in animals that troubles him, but what could come next: "Opening the door to making intermediate species organisms, including humans, is a big cultural step. ... I think there's no natural stopping point when that happens." ■

With reporting by Gretchen Vogel.

PARTICLE PHYSICS

Particle no-show at LHC prompts anxiety

Accelerator enters prime discovery phase, but the clock is ticking

By **Adrian Cho**, in Chicago, Illinois

When physicists working with the world's largest atom smasher, the Large Hadron Collider (LHC), announced last week that much ballyhooed hints of an exotic new particle turned out to be mere statistical fluctuations in the data, many shrugged off the disappointment. Spurious spikes in the data inevitably show up, physicists say, and it's too soon to give up hope for something new and exciting from the LHC, the 27-kilometer-long collider at the European particle physics laboratory, CERN, near Geneva, Switzerland. Yet beneath that equanimity runs a deeper current of anxiety.

The LHC, which started taking data in 2010 but reached high energies only last year, is generating data at an accelerating pace. In physicists' odd units, the collider produced four inverse femtobarns of data last year—about 400 trillion proton-proton

"We could find something by the end of the year. You never know."

Maria Spiropulu, California Institute of Technology

collisions. Coming on great guns, the LHC could have 40 inverse femtobarns by the end of the year. By 2018, it should have 100. But since revealing the famed Higgs boson in 2012, the LHC has failed to unearth a single new particle, and a lack of surprises in the first big batch of high energy data has some physicists concerned. "If we don't find something with 100 inverse femtobarns of data, we should start to worry," says Santiago Folgueras, a physicist at

Purdue University in West Lafayette, Indiana, who works with the Compact Muon Solenoid (CMS), one of four massive particle detectors fed by the LHC.

If the LHC doesn't find anything new by 2018, research there will continue, of course—until 2035, according to the current schedule. However, researchers say, the emphasis would most likely shift from dramatic searches for new, massive particles to more painstaking precision measurements of known particles. Those precision measurements can hint at particles too massive for the LHC to produce directly, but some physicists view the science as less exciting than searches for new particles and say it could lead to attrition at the LHC. “Definitely, I think there would be people who are going to leave,” acknowledges Tiziano Camporesi, a physicist at CERN and spokesperson for the 3000 physicists working with the CMS.

The hints of the new particle, reported this past December by the CMS and its rival detector, A Toroidal LHC Apparatus (ATLAS), raised hopes that a direct search had paid off. It appeared to be five times heavier than the Higgs boson, but like the Higgs it seemed to decay into two photons. But in presentations here at the International Conference on High Energy Physics, the CMS and ATLAS teams reported that the twinned photon signals were flukes, produced by pairs of random,



Time is growing short for new discoveries at the Large Hadron Collider, which is now running at high energies and intensities.

unrelated photons. “I’m not surprised,” says David Charlton, a physicist at the University of Birmingham in the United Kingdom and an ATLAS spokesperson. “It didn’t smell right.”

If real, the particle might have offered a glimpse of physics beyond physicists’ vexing prevailing theory, the standard model. Physicists know that the standard model is incomplete; for example, it does not include gravity. Yet it works frustratingly well, describing every particle blasted into existence with an atom smasher in the past 40 years—including the Higgs boson.

For decades, theorists have dreamed up extensions to the standard model. They generally predict particles weighing a few hundred times as much as a proton and roughly as much as the Higgs: masses

that should be within the LHC’s grasp. For example, some of these particles arise in theories that assume that space has extra spatial dimensions that curl in little loops. Others emerge from a theoretical framework called supersymmetry. That theory, which posits a more massive partner for every particle in the standard model, has excited physicists because it could help solve puzzles such as how the mass of the Higgs itself is stabilized.

Now is the best time to search for such particles with the LHC, physicists say. Hobbled by design flaws, the LHC ran from 2010 to 2013 at half its design energy of 14 tera-electron volts. After repairs, it ran last year near full energy but at low intensity, limiting the chances for statistically significant finds. Now, its beams are also exceeding design intensity. All those data could deliver the longed-for discovery, says Joseph Lykken, a theorist at Fermi National Accelerator Laboratory in Batavia, Illinois, and a member of the CMS team. “I have a bet that we will discover supersymmetry within the first 100 inverse femtobarns of data,” he says.

However, LHC data already have begun to rule out many hypothetical particles over increasing ranges of masses. For example, a key supersymmetric particle called a gluino is now ruled out at masses below about 2000 times that of a proton.

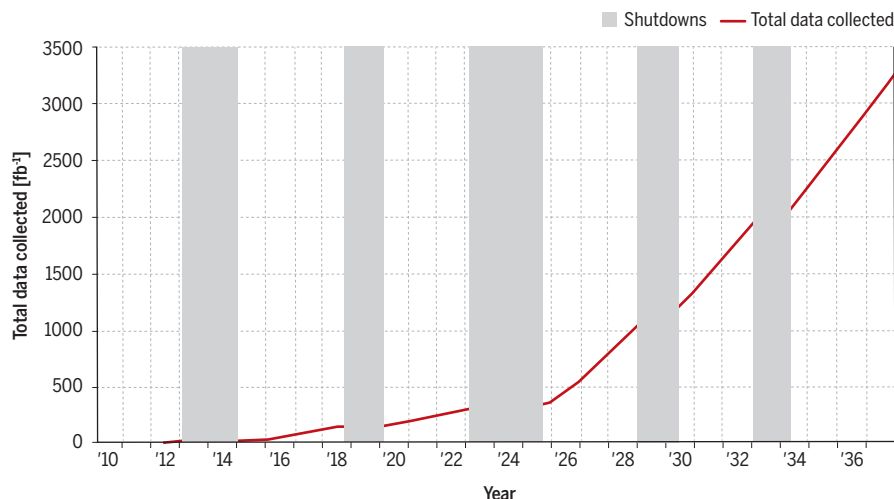
If direct searches for new particles don’t pay off soon, work will shift to precision measurements of the known particles. The idea is to use the LHC’s increasing intensity to create familiar particles such as the Higgs and the top quark in vast numbers, and look for rare anomalies in their decays. That approach could reveal hints of particles too massive to be produced directly. Thanks to quantum uncertainty, these particles flicker in and out of existence in the vacuum, and their ghostly presence could skew the decay patterns of known particles—a less dramatic route to new physics than discovering a new particle.

Some say that shift to precision work could come as soon as the end of 2018. In fact, CERN is already planning for that shift—but not so soon. In 2023, the LHC will undergo a major upgrade to increase the intensity of its beams by a factor of five to seven.

There is still a real possibility that physicists won’t have to face that future, says Maria Spiropulu, a CMS member from the California Institute of Technology in Pasadena. “We could find something by the end of the year,” she says. “You never know.” ■

Smashing success

By the end of 2016, the Large Hadron Collider will have collected nearly twice as much data as it had before its first shutdown in early 2013. A planned upgrade in 2023 would further boost beam intensities and data rates. Those capabilities would better serve the precision measurement of known particles than the exploration at high energies that has excited many physicists.





ECOLOGY

Scientists split on Oregon old-growth forest plan

The northern spotted owl is again the center of controversy

By Warren Cornwall

Two decades after a sweeping government plan curtailed most logging in old-growth federal forests in the Pacific Northwest, the timber wars are flaring again.

The federal Bureau of Land Management's (BLM's) new rules governing 1 million hectares of forest in western Oregon, released on 5 August, have drawn swift criticism from both environmentalists and timber interests. They have also laid bare disagreements among scientists over the best way to restore the northwest's forests from decades of intensive logging and revive the creatures that live there, including the endangered northern spotted owl (*Strix occidentalis caurina*). The stakes are high, because the BLM rules could provide a blueprint for managing all the northwest's vast public forests.

One architect of the original 1994 plan, known as the Northwest Forest Plan, praises BLM's new approach as a much-needed update that incorporates emerging science about forest ecosystems and current conditions in the woods. Another prominent conservation biologist, however, warns the new approach meddles with a plan that's working and whittles away at protections in the name of cutting down more trees. "We should be

celebrating the [original] Northwest Forest Plan. It's not broken," says Dominick DellaSala, president of the nonprofit Geos Institute in Ashland, Oregon, and past president of the Society for Conservation Biology's North American branch.

BLM has been under political pressure for years to boost logging on its lands included in the Northwest Forest Plan. The bureau says its new rules are an attempt to balance conservation with the economic needs of local governments, which under a 1937 law receive half the logging revenue from most of these BLM lands. The new approach is expected to increase logging rates by approximately 37%, to 278 million board feet a year. Seventy-four percent of the land would be managed primarily for ecological benefits with limited logging, an increase of 11%. The rules would also loosen some restrictions on logging near streams or some protected species.

Jerry Franklin, a forest ecologist at the University of Washington, Seattle, was part of the "gang of four," a group of four scientists who helped craft the original plan. He says BLM's changes are an improvement, in part because they embrace logging as a tool for ecological restoration. "I would say in general the way the plan has moved is quite impressive. It's quite progressive and generally reflects the way that science has moved," he says.

Revised federal rules relax some restrictions on logging in the Pacific Northwest.

He still considers logging inappropriate for older, wetter forests filled with trees 80 years old or more. But he has come to believe that it can help drier southern Oregon forests overcrowded by smaller trees left from decades of fire suppression. In wetter parts of the state, he thinks more logging in younger forests can reintroduce the vital brushy landscapes once created by wildfires.

The new BLM plan calls for leaving pockets of forest standing in most logged areas rather than full clearcutting. It would also let logged areas regrow without intensive management, with the aim of allowing more of the brushy lands Franklin wants.

But DellaSala warns the forests are only 20 years into a recovery that could take a century. Rather than turning to logging as a restoration tool, he says land managers should be patient and wait for natural processes, such as forest fires, to do the work. He criticizes BLM's move to shrink by half the forested buffers along fish-bearing streams, to as little as approximately 50 meters. He also worries that the rules permit more logging in some land designated as critical spotted owl habitat even though the bird's numbers have continued spiraling downward. "Every acre of spotted owl habitat is critical. This is just a recipe for extinction," DellaSala says.

Mark Brown, BLM's project manager in Portland, Oregon, says the vast majority of habitat important to spotted owls and marbled murrelets, another federally protected bird, is shielded from logging. The U.S. Fish and Wildlife Service, which is responsible for restoring these species, has endorsed the approach. BLM is trying to protect forests while at the same time meeting legal requirements to generate timber revenue, Brown says. "That's the balance that we've sought here."

But the threat of lawsuits is already in the air. Environmental groups, including the law firm Earthjustice, have echoed DellaSala's critique. Meanwhile, the Association of O&C Counties, which represents 17 counties getting revenue from these BLM lands, filed suit against the heads of BLM and the U.S. Department of the Interior last week, arguing the new plan broke the law by not allowing enough logging.

The outcome of this dispute will set the stage for plans to revise the rules governing the much bigger collection of northwest forests under Forest Service control, predicts K. Norman Johnson, a forestry professor at Oregon State University, Corvallis, who helped craft the 1994 plan. "In some ways the BLM is the Forest Service's β tester of new ideas," he says. "The Forest Service will be watching very closely." ■

FEATURES



FROM BARK TO BEDSIDE

Clinical trials in pets are ramping up, but will they benefit human health?

By David Grimm

Frankie, a 15-year-old brown dachshund with a gray muzzle and tired eyes, rests on a pillow and pink blanket on an exam table at the University of Pennsylvania School of Veterinary Medicine (Penn Vet). A catheter draws blood from her neck into a gray machine the size of a minifridge, which clicks and whirls as it returns clear fluid to her body through another tube. The dog is strapped down by a red leash, but the restraint hardly seems necessary; she looks like she could fall

asleep at any moment. At least until veterinary internist J.D. Foster sticks a thermometer in her butt.

A black mass has engulfed Frankie's bottom-right canine tooth—a melanoma that could eventually metastasize. If her owner had taken her to a traditional vet clinic, the doctor would have likely recommended removing part of her jaw, followed by a strong course of chemotherapy. But here at Penn Vet, Foster and his team are trying a new approach: cleaning Frankie's blood with an experimental polymer that

removes immune system blockers, which may allow her to better fight the cancer. If the treatment works, it probably won't extend Frankie's life—but it could make her last few months a lot more pleasant. It also just might lead to a new way to combat skin cancer in people, Foster says. He scratches Frankie behind the ears as she closes her eyes on the pillow. "You're a good girl," he coos.

Frankie is the third of 13 canines in the study—a clinical trial that's part of a growing push to develop new therapies for people by

Frankie the dachshund takes part in an experimental melanoma trial at the University of Pennsylvania School of Veterinary Medicine.

testing them in sick dogs and cats instead of lab rats or mice. Pets are a better model of human illnesses than rodents, advocates contend: They live in the same environments, sometimes eat the same food, and get many of the same diseases, particularly cancers, that we do. So, the thinking goes, they could hold the key to developing new therapies for humans at a fraction of the normal cost—and potentially yield a trove of new medicines for pets themselves.

“There’s an opportunity for everyone to benefit,” says Amy LeBlanc, who oversees pet clinical trials at the U.S. National Institutes of Health (NIH) as the director of its Comparative Oncology Program in Bethesda, Maryland. The number of such trials is booming, she notes, with hundreds conducted over the past decade.

But others question whether these studies will really have an impact on human health. “It’s a very interesting idea, and it all sounds very nice,” says Larry Baker, an oncologist at the University of Michigan, Ann Arbor, and the former chairman of one of the largest human clinical trial organizations in the United States. “But this field has yet to prove itself.”

PET CLINICAL TRIALS date back to the mid-1970s, when vets tested lymphoma vaccines in household dogs. Later trials tackled other cancers with bone marrow transplants and inhaled chemotherapies—all with the hopes of translating the approaches to humans. But the studies were few and far between, and there was no coordinated effort to bring them into the mainstream.

That began to change about a decade ago. More than half of U.S. households had pets by that point, and owners were spending billions of dollars a year on veterinary care alone. There was also a growing crisis in the traditional drug development pipeline. Typically, researchers first test experimental therapies in laboratory rodents, then in larger animals such as monkeys, before moving to human trials. But the success rates have been abysmal. It can take 16 years to bring a new drug to market, at a cost of up to \$2 billion, and 10,000 promising compounds may lead to only one U.S. Food and Drug Administration (FDA)-approved treatment. Cancer therapies fare

particularly poorly: Only 11% of oncology drugs that work in mice are ever approved for people.

“Drug development is not a sustainable model the way it is now,” LeBlanc says. “It’s too much money and too much time.”

Enter pets. Unlike lab animals, which are highly inbred and raised in environments so sterile they typically only get cancer when researchers obliterate their immune systems and inject them with tumors, dogs and cats are genetically diverse and live in the real, messy world. So it’s no surprise that they get diseases, from cancer to arthritis to muscular dystrophy, that seem to mimic ours. A type of breast cancer in cats has been associated with the same protein—HER2—as one of the most aggressive breast cancers in women, and the bone cancer osteosarcoma is nearly identical—both clinically and genetically—in dogs and people. That may be why a lot of the same drugs help both us and our pets. The most common therapies for human non-Hodgkin

chief science officer with Ethos Veterinary Health, a new Woburn, Massachusetts-based network of veterinary hospitals that plans to conduct clinical trials. “The mouse has prove[n] itself time and time again to be a bad investment. The dog is an alternative to something we know doesn’t work.”

A COUPLE OF FLOORS DOWN from Frankie, a brindle pit bull named Paisley is wagging excitedly in a small exam room. Her owner, Chelsea Burns, holds her tight on a black leash while a vet tech peppers her with questions: How often does she eat? How are her bowel movements? Burns hands over two grocery store bags filled with poop. “She’s an overachiever,” she smiles sheepishly.

Paisley is here for a different clinical trial, one that aims to figure out whether the gut microbiome—the ecosystem of intestinal bacteria linked to everything from allergies to obesity—can give clues to which therapies best treat irritable bowel disease. Today, Paisley will have a physical exam and an ultrasound to confirm she has the condition. Tomorrow, she’ll get a colonoscopy and a biopsy of her intestines. If she qualifies for the trial, she and 50 other dogs will spend 7 weeks on a special diet—and possibly a course of antibiotics—with regular trips to Penn Vet for evaluation.

It all sounds a lot like a human clinical trial, and it is—replete with many of the same challenges. Burns has to sign an eight-page consent form, for one, and some pet clinical trial proponents have suggested adding a “patient advocate” to such studies to ensure that someone other than the emotionally attached owner is looking out for the cat or dog’s best interest—two lines of red tape rat researchers don’t have to cross.

What’s more, owners like Burns consider their pets part of the family (“Paisley’s pretty

much like my child,” she says), so there is a limit to the type of experiments that can be done on them. No veterinarian is going to euthanize someone’s dog at the end of a trial to get a better sense of how a drug worked, for example, even though that could be incredibly informative. “You can try just about anything in the rodent world,” says Dottie Brown, the director of the Veterinary Clinical Investigations Center at Penn Vet. “But these are people’s pets.”

Dogs and cats have other disadvantages. The same things that make them good



Paisley the pit bull gets an ultrasound as part of a clinical trial on irritable bowel disease at the University of Pennsylvania School of Veterinary Medicine.

lymphoma work in dogs, and those that don’t work in dogs don’t work in people.

Researchers can also conduct more detailed studies on pets than they could in the past, thanks to advances in veterinary medicine—from kidney transplants to stem cell treatments—and breakthroughs like the sequencing of the dog and cat genomes. “We can ask much more scientifically rigorous questions that are more likely to intersect with human health,” says Chand Khanna, who founded NIH’s Comparative Oncology Program in 2003 and is now the

models for human disease—more diverse environments and genetics—cause problems when scientists are trying to eliminate variables that could tell them why a drug didn't work. There are also plenty of human diseases that pets don't get, like prostate cancer and Parkinson's. Plus, researchers don't have to spend months recruiting lab rodents, and they can give them a much smaller dose of an experimental drug. That makes pet trials expensive, as does the fact that these studies—like those in people—typically cover the cost of care and procedures, which can run into the thousands or tens of thousands of dollars, especially when dealing with cutting-edge therapies like bone marrow transplants and radiation for brain tumors. "Doing a clinical trial in dogs is 10 times cheaper than doing it in

in animal health studies. Things might be easier if the field had notched some big successes. But Baker, who made his reputation treating children with osteosarcoma, says he hasn't seen many. He points to a couple recent trials that used experimental drugs to treat osteosarcoma that looked promising in dogs but fell short in people. Two other trials testing the antilymphoma drugs GS-9219 and Zydelig proved safe and effective in dogs, but toxic to humans. "We've been so desperate for new drugs, especially for kids, we want to see progress," Baker says. "But that doesn't mean being blinded by our optimism."

LeBlanc says there are signs that pet clinical trials can pay off. In 2013, for example, FDA accelerated its approval of a drug called ibrutinib for lymphoma in people

Academies of Sciences, Engineering, and Medicine, sparked new collaborations, and removed a major roadblock. The pharmaceutical industry had worried that if one of its drugs caused problems in a pet study, FDA would never approve it for people. But at the meeting an FDA representative said that was not the case. "That was a huge win for the field," Khanna says.

Since then, NIH has become more interested in these trials, too. In April, the National Cancer Institute announced that it would fund partnerships between veterinarians, physicians, and scientists to better understand the molecular biology of dog cancer. NIH also inspired the American Veterinary Medical Association to create its own version of ClinicalTrials.gov (a worldwide database of human clinical trials), which went online in June. And next year, the agency's Comparative Oncology Trials Consortium—a network of 22 North American academic centers—will begin releasing data from its largest pet clinical trial to date, a 160-dog study of canine osteosarcoma, which LeBlanc is overseeing. "We want to get better data into everyone's hands," she says.

The field is also hoping for its first big translational success. In 2015, New Haven, Connecticut-based Kolltan Pharmaceuticals announced that its experimental antibody drug KTN0158 dramatically shrank a common skin tumor in dogs, prompting it to begin clinical trials in humans. "The dog trial had a dramatic impact on our strategy to develop this product for people," says the company's president, Jerry McMahon.

He believes the main role of pet clinical trials in the future will be to help pharmaceutical companies like his minimize the risk of drug development. Instead of putting every drug that works in mice into the human pipeline, they can focus on the ones that help sick cats and dogs. "It makes the investment in human drug development more promising," he says. "It takes a while for things like pet clinical trials to become part of mainstream drug development, but I think they could be a powerful tool for the future."

Of course, there's always the possibility that pet clinical trials will never translate to people, and that they'll just help veterinarians develop better drugs for dogs and cats. But many advocates don't see that as a bad thing. "If we save these dogs, it has an impact on every single family that owns a dog," says Matthew Breen, a geneticist at North Carolina State University in Raleigh who has been part of more than a dozen pet studies. "When I get involved in these trials, it's about helping the family. If we're helping the human or the dog, is there really any difference?" ■

Sit. Stay. Heal. A sampling of pet clinical trials

Researchers are conducting hundreds of clinical trials on dogs and cats across the world. Here are a few that may have an impact on human health.

TRIAL	PURPOSE	LOCATION
Minibeam radiation for brain tumors	Testing whether small, parallel beams of radiation can treat malignant brain tumors in dogs.	U. of Saskatchewan, Saskatoon, Canada
GammaCore VET for seizures	Testing whether a handheld electronic nerve stimulator can treat epilepsy in dogs.	U. of Georgia, Athens
HER2/neu breast cancer vaccine	Testing an experimental vaccine against a type of breast cancer in cats that is also one of the most aggressive in women.	Veterinary Oncology Services, Middletown, New York
Stem cells for dry eyes	Testing whether an injection of stem cells from a dog's own fat can restore the eye's ability to tear.	Purdue U., West Lafayette, Indiana
Photodynamic therapy for lung tumors	Combining a drug with laser light to kill lung tumor cells in dogs.	Colorado State U., Fort Collins
Embolization for prostate cancer	Testing whether blocking blood vessels in the prostate gland can treat prostate cancer in dogs.	U. of California, Davis
Canine osteoarthritis pain	Testing two experimental anti-inflammatory medications.	U. of Minnesota, Twin Cities

humans," Brown says, "but 10 times more expensive than doing it in rats."

Those are all big reasons it's hard to get funding for these trials, LeBlanc says, especially from pharmaceutical companies. "They want to get their drug to market as soon as they can."

The government has also been loath to fund pet clinical trials because scientists know a ton more about rats and mice than they do dogs and cats, LeBlanc says. "NIH gets really uncomfortable with models that haven't been vetted in the same way as mouse models." That has forced veterinarians to scrounge money from their institutions or from pet-centric organizations, most prominently the Denver-based Morris Animal Foundation, a nonprofit that claims to have invested more than \$100 million

after it showed promise in canine trials. And because dogs are easier to operate on than rodents, they have helped optimize ways to excise bone tumors in children without removing limbs.

Still, LeBlanc admits that the field has a way to go. "We don't have a slam-dunk transition to better human health," she says. "But I think that data is coming."

THERE ARE HINTS that pet clinical trials may be edging into the mainstream. Last summer, nearly 200 veterinarians and physicians met for the first time with funders, government regulators, and representatives from the pharmaceutical industry in Washington, D.C., to discuss how these studies could better contribute to human health. The workshop, sponsored by the U.S. National



Nine months after a spill, a stick stirred up oily water in a canal near Cuninico, Peru.

OIL IN THE FOREST

A rash of spills in the Amazon has scientists and villagers scrambling to understand oil's effects on the ecosystem

By **Barbara Fraser**, in Cuninico, Peru

Spring should have been a good time for fishing in the network of dark, tree-lined lakes around this small Kukama Indian village in the northern Peruvian Amazon. Seasonal floodwaters that draw river fish into the lakes to spawn have traditionally meant a bonanza for fishermen here. But this year was different.

"There are no fish in the lakes," says village President Galo Vásquez, a small, soft-

spoken man with a perpetually worried expression. "We have to go farther upriver, sometimes a day or two away."

A late start for the rainy season could be to blame. But Vásquez and his neighbors also suspect lingering effects from a 2014 oil spill along a pipeline that carries crude 845 kilometers from Amazonian oil fields to the Pacific coast. Traces of oil still occasionally float past the village, even after 2 years of cleanup efforts in the area.

Sitting down to breakfast at a wooden

table in an open-sided house typical of the Amazon, as rain pattered on the palm-thatch roof, Vásquez and his neighbors wondered aloud whether the pollution drove the fish away and whether the few fish they did catch were safe to eat. A recent toxicology study found heavy metals in villagers' blood and urine, but did not pinpoint a source, raising more questions than it answered.

The dense green canopy of the rainforest in eastern Peru and Ecuador hides substantial reserves of oil, and over the

last 4 decades a web of wells and pipelines has spread through the forest. And where there is oil, there are spills. In the past few years, a spate of oil spills along Peru's main oil pipeline, including three this year alone, has fouled the water and land of dozens of indigenous villages. The spills have sparked a lawsuit by affected communities—and fledgling efforts by scientists to understand how oil affects the Amazon ecosystem.

By global standards, the Amazon spills are small. The 2000 barrels that leaked from the pipeline here are a drop compared with the more than 3 million barrels that spewed out of the sea floor when the *Deepwater Horizon* drilling rig exploded in the Gulf of Mexico in 2010. But oil may affect the complex Amazon ecosystem in unexpected ways, so that lessons learned in marine spills may not apply. “Where biodiversity is incredibly high,” as in the Amazon, each species may react differently, says physical geographer Paúl Arellano of Yachay Tech in Urcuquí, Ecuador. And because local people rely on fishing and on river water for drinking, the human impacts of oil pollution could be especially severe.

A handful of studies are beginning to trace toxic components of oil as they wend their way through the Amazonian ecosystem, into the river fish and possibly into the people who eat them. Still, “of all we know about the fate of crude oil in water in general, only 2% or 3% of the information [comes from] freshwater systems,” says biologist Adalberto Val of Brazil's National Institute of Amazonian Research in Manaus. “We are just starting to [learn about] this in the Amazon.”

A FISHERMAN FROM CUNINICO

raised the alarm on 30 June 2014 when he saw an oil slick and dead fish in a stream that flows from a canal containing the submerged pipeline to the Cuninico River. Workers from Petroperú, the state-owned oil company that operates the pipeline, arrived quickly and hired men from the community to search for the rupture. The men worked for days in chest-high, oily water, at first without the protective gear standard in the United States. Later, the company hired a cleanup firm, and by October 2014, most visible oil was gone.

In December 2014, however, the rainy season began. The river rose, and booms meant to corral oil in the pipeline canal floated free of their moorings. Oily water overflowed into the forest and the Cuninico River.

Petroperú workers returned last fall. But in March, nearly 2 years after the spill, a village fishing near the mouth of the Cuninico River caught fish that smelled of gasoline, Vásquez says. Even now, stirring the bot-

tom of the canal with a long pole releases oily bubbles.

Amazon itself. And fish in both types of rivers have adapted to fluctuations in oxygen levels, which rise during the day and during flood season and fall at night and in the dry season. Some fish actually come to the surface to breathe air, as do river dolphins; others seek oxygen-rich shallow water.

These adaptations can become a liability in an oil spill, as a handful of studies show. Because slicks typically form at the surface, river dolphins and the air-breathing fish *Arapaima gigas*, known as pirarucu in Brazil and paiche in Peru, ingest oil when they breathe, Val says. Fish also are exposed to dissolved hydrocarbon compounds that are more toxic near the surface, thanks to chemical reactions catalyzed by ultraviolet light, according to a 2002 study.

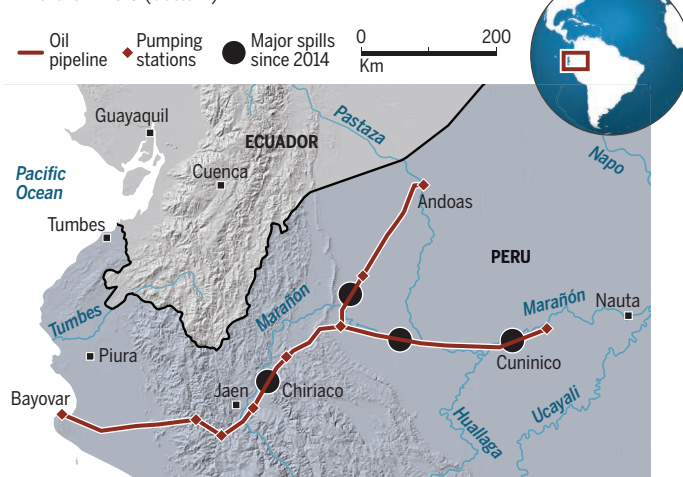
Cleanup crews sometimes apply dispersants—detergent-like chemicals famously used in the *Deepwater Horizon* spill—to break oil into smaller particles so that microbes can biodegrade it more quickly. In laboratory tests, however, Val and his colleagues found that a combination of oil and a dispersant can be more toxic to Amazonian fish than oil alone, as they reported in 2010. In another paper just a few weeks ago, they reported that oil and a common dispersant appear to be more toxic in “black water” than in groundwater, causing liver and DNA damage in the fish.

The fish that survive a spill—potential catch for locals—may be contaminated. After spills early this year, fish farther up the nearby Marañón River contained levels of cadmium and lead above the limits set by Peru's Ministry of Health. No one can say whether some species are safe to eat, however, and it's not clear whether the metals came from oil or other sources such as mining.

WHATEVER THEIR SOURCE, heavy metals may be making their way into people. At village leaders' insistence, officials from Peru's National Institute of Health in January analyzed blood and urine samples from 126 people here and in nearby San Pedro. Cadmium and mercury were above the hazard level in two-thirds of people tested, physician Jonh Astete of the National Institute of Health's occupational and environmental health of-

Oil across the Amazon

A pipeline ferrying crude oil 845 kilometers from Amazonian fields across the Andes Mountains to the Pacific coast has suffered repeated spills. In January, a spill fouled a stream near the town of Chiriaco in northern Peru (bottom).



tom of the canal with a long pole releases oily bubbles.

The oil has entered a complex ecosystem, where floodwaters rise and fall, water chemistry varies from river to river, and fish have diverse adaptations that can affect how they absorb pollution from the water and sediments. For example, fish in acidic, organic-rich “black water” rivers like the small Cuninico absorb minerals differently than fish in “white water” rivers like the

fice told villagers here in June. And about 75% of children had detectable levels of lead in their blood, although only two exceeded Peru's threshold of 10 micrograms/deciliter (mg/dl) for children. (In the United States, the threshold is 5 mg/dl, although experts agree that no level of lead is actually safe for children.)

The findings mirror those of a 2005 study of villagers near legacy oil fields developed decades ago along Peru's Corrientes River. Early on when rules were lax, spilled oil—plus the hot, salty, metals-laden water pumped out with it—flowed into streams and lakes and was not cleaned up. (By law, such fluids are now reinjected underground.) Contaminated sites often went unremediated or were covered over, a practice that has led to lawsuits.

Along the Corrientes, two-thirds of the 75 children tested had blood lead levels of between 10 and 45 mg/dl, a level that experts agree can affect children's cognitive and motor development. About 80% of the 124 adults tested had blood lead levels between 10 and 20 mg/dl, and nearly everyone had blood cadmium levels higher than Peru's reference level.

To work out whether such metals really come from spilled oil, researchers say they need to track oil through sediments and flooded forests. Monitoring polyaromatic hydrocarbons—among the most toxic substances in oil—in fish tissue is also a priority. Astete says a new study of about 2500 people in villages along four rivers in the region, including the Marañón and Corrientes, could help pinpoint the origin of the heavy metals—whether from spills, industrial activities such as oil drilling or mine runoff, burning of trash near homes, or natural sources.

As part of the study, a team from his agency set up a makeshift laboratory here in early June, took blood and urine samples from members of more than 30 families, and filled out detailed questionnaires. They also collected samples of food the families had prepared, including fish, cassava, and bananas. They even bagged a small caiman that a hunter had caught along the Cuninico River, but had decided not to eat because of possible contamination. Samples from both people and food will be tested for arsenic, metals, and hydrocarbons.

Researchers returned in July for sediment, water, soil, and air samples. They hope that comparisons of contamination in

people, food, and the environment will reveal the pollution's trail.

While samples are being tested, Astete urged Cuninico residents to avoid fish-eating fish, which are more likely to concentrate some pollutants in their bodies, as well as bottom feeders that could absorb pollutants from sediments. But locals say these limits eliminate some of the most commonly eaten fish.

Such public health measures are harder to target when the contamination comes from “legacy oil,” dumped years ago and often hidden. Arellano says that during field research he stumbled on places where workers had dumped oil into ponds more than a decade earlier and then covered them up;

Arellano is now using satellite data to try to locate similarly stressed trees. His method could be useful for estimating legacy contamination, if he can distinguish the photosynthesis reduction caused by oil pollution from that due to other possible causes, says biologist Gregory Asner of Stanford University in Palo Alto, California, who uses spectrometry to map Amazon biodiversity.

Remote sensing may also help agencies and communities monitor spills from future operations. Low oil prices have lulled Amazonian operations, but new projects are planned, some in or near protected areas, including Ecuador's biodiverse Yasuni National Park and Peru's Pacaya Samiria



Amazonian fishers like these on Monkey's Lake in Brazil rely on the giant pirarucu, called paiche in Peru. But the air-breathing fish rise to the surface often, exposing them to contamination if an oil spill occurs.

shrubs and trees gradually grew back. “You can be walking along and have no idea that you're on top of a pool of oil, but if you dig down a little, you find crude.”

Remote sensing may be able to reveal these hidden pools, Arellano says. He compared the leaves of canopy trees in an abandoned spill site with those in uncontaminated ecosystems. Leaves in the spill sites contained less chlorophyll, which subtly changes their color, he reported last year. The leaves also contain more water, suggesting that their pores (or stomata) had closed up, reducing photosynthesis. You can't spot these trees by sight, but satellite imagery can detect the color change starting about 3 weeks after pollution occurs and for decades afterward, Arellano says.

National Reserve, a huge expanse of wetlands and seasonally flooded forest.

Back in Cuninico, the rainy season has ended and the river is subsiding. It will be another year before villagers know whether this year's poor fishing was likely due to delayed rains or pollution. Community leaders have filed a lawsuit against government agencies they say failed to protect them. Any compensation will help, but villagers wonder whether their lives will ever be the same.

“People are still worried,” says César Mozombite, a member of the village's leadership council. “They're not catching fish.” ■

Barbara Fraser is a freelance writer based in Lima.

INSIGHTS

PERSPECTIVES

ASTRONOMY

Making sense of the exoplanet zoo

The diversity of substellar objects may call for a rethinking of our labeling conventions

By **Rebecca Oppenheimer**

The regions around stars on length and mass scales similar to that of our own solar system are relatively new to human exploration. In the two decades since “substellar objects,” things less massive than stars, were discovered orbiting stars other than the Sun, the single most certain statement about them has been “expect the unexpected.” On page 673 of this issue, Wagner *et al.* (1) reinforce that statement by reporting on the detection of a substellar object orbiting within a triple-star system.

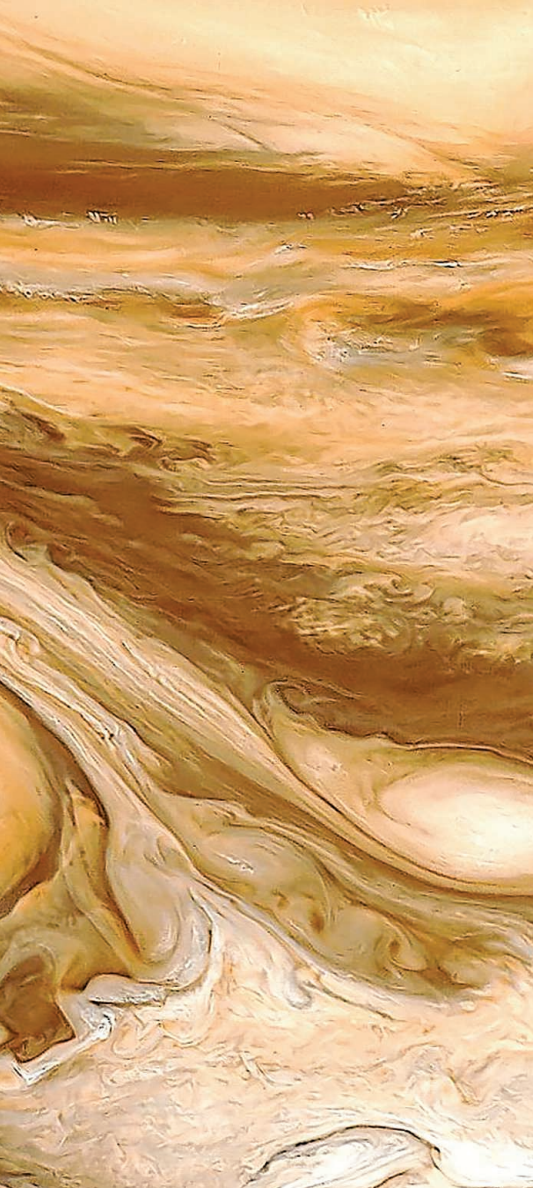
Although a number of other substellar objects have been observed in triple-star sys-

tems (2–5), the one reported by Wagner *et al.* is particularly interesting because it must be influenced by the gravitational pull of the other two stars, causing an improbable orbit. In fact, this small object, somewhat more massive than Jupiter, yet very young (~16 million years), could well be in an unstable orbit. It could be thrown out into space to drift alone. Many such solitary objects, some perhaps rejected from their natal solar systems, are being discovered routinely (6). All are different from each other, straining current classification schemes.

Another star, one of thousands now discovered, that exhibits the “unexpected” and draws into question what constitutes a solar system is HD 41004. It is somewhat smaller

than the Sun, with an object 2.5 times as massive as Jupiter on an orbit slightly more than Earth’s about the Sun. In addition, another star orbiting HD 41004, at the equivalent of Uranus’s orbit, has a substellar object orbiting it with about 20 times the mass of Jupiter (4, 5). So, is our labeling of HD 41004 as a “solar system” accurate even though it contains two stars, an exoplanet, and a brown dwarf (an object intermediate between a star and a planet)? Also, is that brown dwarf actually a moon of the smaller star? The more general question to ask our-

Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192, USA. Email: roppenheimer@amnh.org

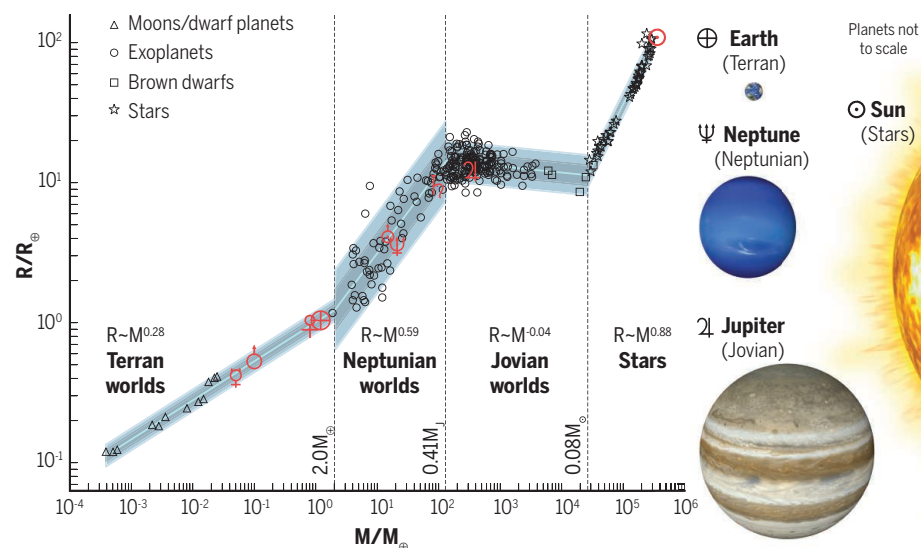


Jupiter's Great Red Spot. A snapshot of the swirling dynamics of the Jovian planet's atmosphere. But, what characteristics define other substellar objects?

selves is what these words actually mean.

Categorizing is an age-old practice in scientific thought. Labels imply meaning and assign properties to things, so that they can be discussed efficiently. However, after 22 years of working on substellar objects, I suspect that these categories may have lost their utility in advancing knowledge (7). Labels can become obfuscating jargon.

To study in isolation what we call stars, brown dwarfs, planets, moons, or even dwarf planets implies that they are intrinsically unrelated entities. Perhaps this is understandable: As searches for substellar objects continue, the enormous diversity in salient properties appears as one examines objects of smaller mass. Every single object requires its own explanation through many parameters (8). Age, chemical composition, mass, temperature, irradiation from a nearby star—all of these factors, and oth-



Cataloging categories. Mass-radius relation for objects generally referred to as planets, brown dwarfs, and stars [reproduced from (10)]. Our current classification schemes using terms such as “exoplanet,” “brown dwarf,” or “star” may be of limited value for scientific understanding. Radius is represented in units of Earth radii, and mass in units of Earth mass. The historical signs (red) represent the various planets of our solar system. The apparent scarcity of “brown dwarfs” (black squares) is simply because few have well-measured masses and radii, although thousands have been found.

ers, determine what an object looks like. Some young substellar objects even look like old stars (1, 5, 6, 8). In contrast, stars, if one can measure their brightness and distance, reveal their mass, and how they live their lives, with particular certainty (9). With fascinating discoveries, such as Wagner *et al.*'s new “planet” in a triple-star system, and the thousands of objects intermediate between it and stars, what we know is that they consistently fail to conform to the stellar classification system intrinsic to the history of astrophysics (8). In such a confusing situation, the best we can do is to rethink the basic assumptions.

Some scientists are doing just that, and perhaps revealing relationships between these seemingly disparate substellar objects. The most fundamental parameters that we can measure for substellar objects are their mass and radius. Chemical composition, temperature, and orbits are also measurable, but individual objects can evolve (and do) with respect to these quantities. Mass and radius are more constant (except radius at very young ages). Both are extremely difficult to measure, but we are now beginning to amass precise measurements for hundreds of objects.

From the accurately known masses and radii of substellar objects, it is clear that there are different groupings (see the figure) (10). Massive objects that fuse hydrogen (stars) have a different relation between mass and radius than lower-mass objects that cannot sustain long-term nuclear fusion. The lower the mass of those objects, the larger they are (analogous to Jupiter, including what are

called brown dwarfs and exoplanets). At even lower masses, worlds like Neptune follow a different relation and are smaller at smaller masses. At the lowest masses, solid objects like Earth and the Moon behave like objects we tend to know in everyday life—they are hard, smaller, and less massive. Chen and Kipping (10) propose a new nomenclature based on these observed measurements: Jovian, Neptunian, and Terran worlds. Whether this scheme is useful will certainly be debated, but it is a fresh alternative to the confusing terms in use today.

Perhaps it is too early to define classes of objects inhabiting our universe's vast zoo of diverse solar systems. To do so may obscure their commonalities and differences, urging overspecialization in the study of objects assumed to be unrelated because of thought-constraining labels. The new system of objects described by Wagner *et al.* serves as a reminder that the universe is full of unexpected phenomena. ■

REFERENCES AND NOTES

1. K. Wagner *et al.*, *Science* **353**, 673 (2016).
2. B. Macintosh *et al.*, *Science* **350**, 64 (2015).
3. T. Dupuy *et al.*, *Astrophys. J.* **817**, 80 (2016).
4. S. Zucker *et al.*, *Astron. Astrophys.* **426**, 695 (2004).
5. G. Chauvin *et al.*, *Astron. Astrophys.* **528**, A8 (2011).
6. There are now many examples of “free-floating planets.” See, most recently, (11).
7. B. R. Oppenheimer, in *50 Years of Brown Dwarfs: From Prediction to Discovery to Forefront of Research*, V. Joergens, Ed. (Springer, 2014), pp. 81–111.
8. J. K. Faherty *et al.*, <http://arxiv.org/abs/1605.07927> (2016).
9. H. Vogt, *Astron. Nachr.* **226**, 301 (1925).
10. J. Chen, D. M. Kipping, <http://arxiv.org/abs/1603.08614> (2016).
11. S. Maer *et al.*, *Mon. Not. R. Astron. Soc.* **461**, L107 (2016).

10.1126/science.aah4097

STRUCTURAL BIOLOGY

Twists and turns in gating ion channels with voltage

The EAG potassium channel may use an alternative voltage-dependent gating mechanism

By **Gilman E. S. Toombes** and
Kenton J. Swartz

The discovery of mutations that make flies shake or dance was a major breakthrough in neuroscience (1) because it led to the identification of genes for an array of ion channels that are selectively permeable to potassium ions. The gene mutated in the dancing flies, christened “Ether-a-go-go” (EAG) because the flies move like go-go dancers in response to ether (2), encodes the founding member of the KCNH family of voltage-activated potassium (Kv) channels, which play critical roles in regulating excitability, irregular heartbeat (cardiac long QT syndrome), epilepsy, and cancer (3). On page 664 of this issue, Whicher and MacKinnon (4) solve the structure of the EAG (KCNH1) Kv channel bound to calmodulin (CaM) and discover an unanticipated mechanism by which voltage controls the process of channel opening and closing.

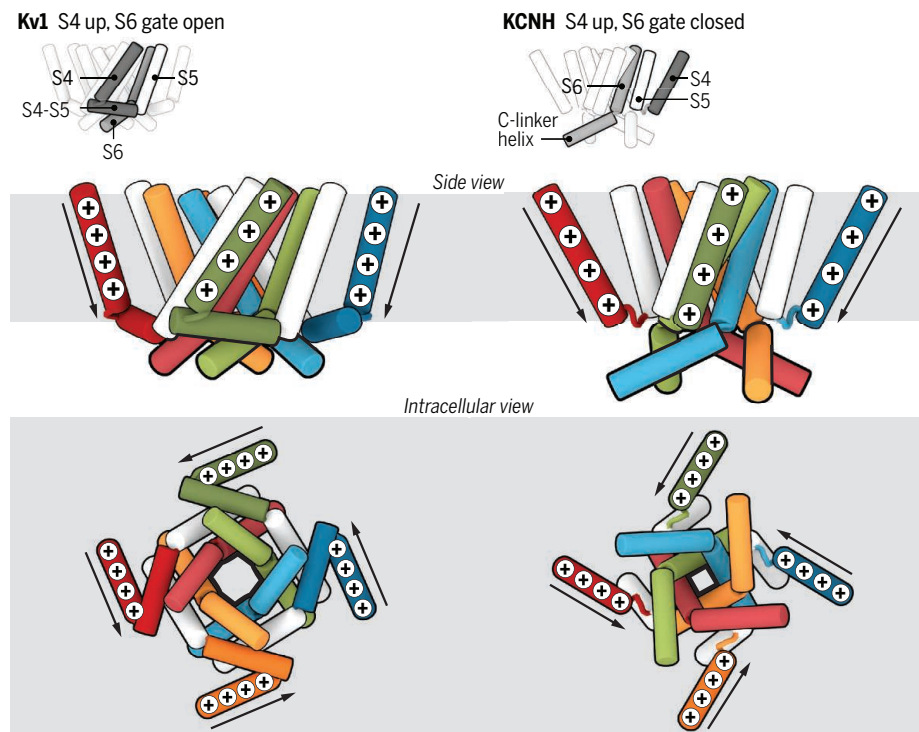
Like other Kv channels (5), the pore domain in KCNH channels is formed by the tetrameric arrangement of S5-S6 helices, and the surrounding voltage-sensing domains are formed by S1-S4 helices (see the figure). This KCNH1 structure gives us the first glimpse of a Kv channel captured in a closed state, where the pore-lining S6 helices converge at the intracellular side of the membrane to obstruct ion permeation.

Unlike other Kv channels, the S6 helices in KCNH1 extend into cytosolic helix-turn-helix motifs termed C-linkers. These motifs are also present in the closely related cyclic nucleotide-gated (CNG) channels and hyperpolarization-activated cyclic nucleotide-sensitive (HCN) channels, where they play a critical role coupling conformational changes of carboxyl-terminal cyclic nucleotide binding domains (CNBDs) to the pore (6). Although KCNH channels are not directly modulated by cyclic nucleotides, their gating properties are tuned by a comparable

cyclic nucleotide binding homology domain (CNBHD) in their carboxyl termini as well as a Per-ARNT-Sim (PAS) domain in their amino termini (7, 8). Compatible with their effects on gating, the CNBHDs are docked under the C-linkers, whereas the more peripheral PAS domains are positioned where their critical amino termini could interact with the S1-S4 voltage-sensing domains. In addition, the structure provides insight into how Ca^{2+} -dependent binding of CaM inhibits KCNH1 channels (9). The authors propose that CaM clamps together the CNBHD and PAS domains from neighboring subunits, thereby rotating the C-linker domains and twisting the S6 helix gate

closed. The most astonishing feature of the KCNH1 structure is the novel coupling between the voltage-sensing and pore domains. In existing structures of Kv channels, as well as voltage-activated sodium and calcium channels (5, 10, 11), each S1-S4 voltage-sensing domain is adjacent to the S5-S6 helices from the neighboring subunit and is connected to its displaced S5-S6 helices via a helical S4-S5 linker running parallel to the intracellular side of the membrane. As a result of this “domain-swapped” arrangement, the S4-S5 linkers encircle and directly interact with the S6 helices, leading to the proposal that they act as levers that enable voltage-driven inward movements of the arginine-rich S4 helices to close the S6 channel gate (12).

By contrast, domain swapping is not seen in KCNH1, where each S1-S4 domain is directly connected to the nearest S5-S6 pore helices via an S4-S5 loop that is too short to interact extensively with the S6 helix. Moreover, Whicher and MacKinnon reason that the voltage-sensing domains are likely to be in an activated state, revealing that CaM binding unexpectedly decouples the voltage-sensing and pore domains. If those



To swap or not to swap? Coupling between positively charged S4 voltage-sensing helices and the intracellular gate within the pore domain is shown for Kv1.2 and KCNH channels. In the Kv1.2 channel, the helical S4-S5 linkers would relay inward movements of the S4 helix to the pore-lining S6 helix in the same subunit, closing the pore. KCNH1, by contrast, has no S4-S5 helices, but inward movement of each S4 helix into the resting state should allow it to interact directly with the C-linker to stabilize a closed pore conformation similar to this CaM-bound state. Top cartoons are side-views within the membrane, and bottom cartoons are from inside the cell. The S4, S4-S5, S6, and C-linker helices from each subunit are shown with a different color. For clarity, only the first helix of each C-linker is shown, and S5 helices are shown in white.

Molecular Physiology and Biophysics Section, Porter Neuroscience Research Center, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA. Email: kenton.swartz@nih.gov

domains can move independently and KCNH1 does not contain a conventional S4-S5 linker, how do conformational changes in S4 trigger opening and closing of the S6 gate? Although a definitive answer will require more experiments, Whicher and MacKinnon propose that at negative membrane voltages, inward movement of each S4 helix would allow it to directly interact with the C-linker helix of the adjacent subunit and thereby stabilize a closed conformation resembling the CaM-inhibited state.

These new structural relations can explain the surprising discovery that voltage-dependent gating in KCNH1 and KCNH2 channels does not require a covalent linkage between S4 and S5 (13), and provide a framework for exploring how inward motions of the S4 helix open rather than close the pore in HCN channels. Indeed, the findings from an earlier investigation of state-dependent interactions between the inner end of S5 and the C-linker in HCN channels are compatible with the new structure of KCNH1, but not that of conventional Kv channels (14).

It is truly remarkable how this single new structure of a voltage-activated ion channel sparks so many fresh ideas about how voltage, nucleotides, and regulatory proteins can control the activity of this singular and diverse family of ion channel proteins. The structure of KCNH1 reported by Whicher and MacKinnon will be a windfall for scientists studying members of the KCNH, CNG, and HCN families of ion channels, and may help us to understand why the related hERG (KCNH2) channel interacts with drugs so often that the pharmaceutical industry must screen for this off-target interaction early in drug development (15). ■

REFERENCES

1. T. W. Kaplan, *Genetics* **61**, 399 (1969).
2. J. Warmke, R. Drysdale, B. Ganetzky, *Science* **252**, 1560 (1991).
3. J. H. Morais-Cabral, G. A. Robertson, *J. Mol. Biol.* **427**, 67 (2015).
4. J. R. Whicher, R. MacKinnon, *Science* **353**, 664 (2016).
5. S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 897 (2005).
6. K. B. Craven, W. N. Zagotta, *Annu. Rev. Physiol.* **68**, 375 (2006).
7. Y. Haitin, A. E. Carlson, W. N. Zagotta, *Nature* **501**, 444 (2013).
8. E. C. Gianulis, Q. Liu, M. C. Trudeau, *J. Gen. Physiol.* **142**, 351 (2013).
9. R. Schonherr, K. Lober, S. H. Heinemann, *EMBO J.* **19**, 3263 (2000).
10. J. Payandeh, T. Scheuer, N. Zheng, W. A. Catterall, *Nature* **475**, 353 (2011).
11. J. Wu *et al.*, *Science* **350**, aad2395 (2015).
12. S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 903 (2005).
13. E. Lorinczi *et al.*, *Nat. Commun.* **6**, 6672 (2015).
14. D. C. Kwan, D. L. Prole, G. Yellen, *J. Gen. Physiol.* **140**, 279 (2012).
15. M. C. Sanguinetti, M. Tristani-Firouzi, *Nature* **440**, 463 (2006).

10.1126/science.aah4194

NEURODEGENERATION

One target for amyotrophic lateral sclerosis therapy?

Targeting a single protein reduces both toxic repeat RNAs and proteins

By Sarah Mizielinska and Adrian M. Isaacs

Repeat expansion mutations cause a range of developmental, neurodegenerative, and neuromuscular disorders. The repeat sequences generally comprise a 3- to 6-base pair repeat unit that expands above a critical threshold, leading to disease. Expanded repeats cause disease via a range of mechanisms, including loss of function of the repeat-containing protein and production of toxic repeat RNAs and proteins, making the disorders difficult to treat. In 2011, a hexanucleotide repeat expansion in the *C9orf72* gene was identified as the most common cause of frontotemporal dementia and amyotrophic lateral sclerosis (termed c9FTD/ALS) (1, 2). On page 708 of this issue, Kramer *et al.* (3) report that targeting a single factor, Spt4, reduced production of *C9orf72* repeat expansion-associated RNA and protein, and ameliorated neurodegeneration in model systems.

Kramer *et al.*'s use of a single factor to reduce multiple repeat-associated pathologies is notable in the light of two unexpected features of repeat expansions. One is that repeat expansions are transcribed in both the antisense and sense direction (4). The other is that repeat-associated non-ATG (RAN) translation occurs, in which repeat expansions mediate their own translation into proteins (5). As no ATG start codon is required, RAN translation can occur in all six sense and antisense frames. A major therapeutic challenge is to target the wide range of potentially toxic RNA and protein species that are produced.

The yeast Spt4 (human homolog SUPT4H1) is a small, evolutionarily conserved zinc finger protein that forms a complex with Spt5. The Spt4-Spt5 complex binds to RNA polymerase II and regulates transcriptional elongation. Deletion of *Spt4* in yeast was shown to reduce the transcription of expanded CAG, CTG, and CAA repeats, but had little effect on short repeats (6). Similar effects were observed with CAG

repeats (which cause Huntington's disease) in cultured mouse neurons. In addition, depletion of *Supt4h* in two different mouse models of Huntington's disease (7) selectively reduced transcription of the repeat expansion allele while leaving transcription of the normal allele unaffected. This decreased mutant, but not wild-type, protein production and aggregation, delayed onset of motor phenotypes, and prolonged life span.

Kramer *et al.* now extend this work to *C9orf72* GGGGCC repeat expansions. In c9FTD/ALS, both sense and antisense repeat RNA transcripts form aggregates, termed RNA foci, in patient brains. RNA

"A major therapeutic challenge is to target the wide range of potentially toxic RNA and protein species that are produced."

foci exert toxicity in other repeat expansion diseases by sequestering essential RNA-binding proteins (8). Using yeast models expressing either expanded sense or antisense *C9orf72* repeats, Kramer *et al.* found that *Spt4* depletion decreased the number of both sense and antisense repeat transcripts and RNA foci.

C9orf72 RAN translation leads to the production of five dipeptide repeat proteins that can cause neurodegeneration in model systems (9). Kramer *et al.* showed that production of one of them, poly(glycine-proline), was substantially reduced by *Spt4* depletion in yeast and worm *C9orf72* models, as would be expected if less repeat RNA was available for translation. Reducing the expression of *Spt4* also improved survival in *C9orf72* worm and fruit fly models, indicating a reduction of toxic repeat species.

To study endogenous expanded repeats, Kramer *et al.* used human c9ALS patient fibroblasts. Reducing the expression of either human *SUPT4H1* or its binding part-

Department of Neurodegenerative Disease, UCL Institute of Neurology, London WC1N 3BG, UK. Email: a.isaacs@ucl.ac.uk

ner *SUPT5H* (homolog of *Spt5*), or both, reduced *C9orf72* sense and antisense RNA foci and poly(glycine-proline) levels. Additionally, reducing *SUPT4H1* expression in c9ALS patient-induced pluripotent stem cell-derived cortical neurons reduced the amount of *C9orf72* transcripts and poly(glycine-proline). Thus, targeting human *SUPT4H1* and *SUPT5H* can effectively reduce multiple key c9FTD/ALS pathologies (see the figure).

One concern for *SUPT4H1* as a therapeutic target is that it may regulate other nonmutated genes. Deletion of *Spt4* in yeast changed the regulation of 149 genes compared to controls (5). In the study of Kramer *et al.*, 95% depletion of *SUPT4H1* in human fibroblasts altered the expression of 301 genes. Of note, deletion of one copy of *Supt4h* did not exhibit any overt phenotype in mice up to 18 months of age, but deletion of both copies is embryonic lethal (6). Thus, the degree of *SUPT4H1* depletion will be critical for effective therapy development.

One of the most advanced potential therapeutics for repeat expansion disorders, including Huntington's disease and c9FTD/ALS, are antisense oligonucleotides that specifically target the mutant expanded allele. Compared to mutant gene-specific antisense oligonucleotides, potential ad-

vantages of the *SUPT4H1*-targeting strategy are its wider applicability and the reduction of both sense and antisense transcripts. However, the relative role of antisense RNA and protein species in disease pathogenesis is currently unclear, so targeting sense repeats may still have a beneficial effect. In this context, Kramer *et al.* show that an antisense oligonucleotide targeting the sense *C9orf72* strand almost completely reduced poly(glycine-proline) levels in patient fibroblasts. In addition, antisense oligonucleotides that specifically target the gene of interest may have fewer off-target effects. As the authors suggest, an exciting possibility is the development of an antisense oligonucleotide targeting *SUPT4H1*, particularly because this may have broad potential for repeat expansion diseases. ■

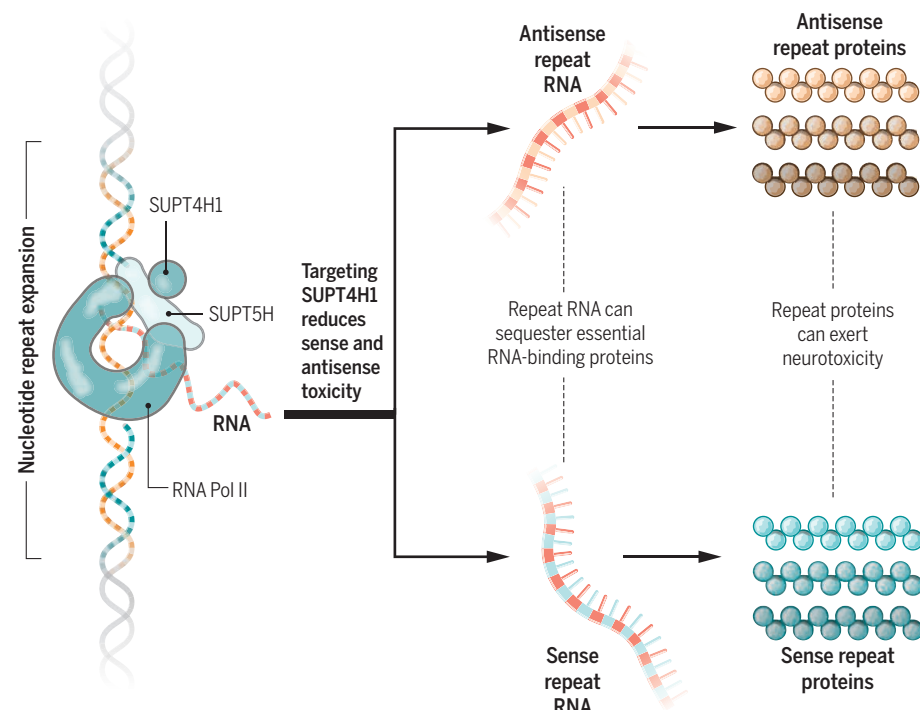
REFERENCES

1. M. Dejesus-Hernandez *et al.*, *Neuron* **72**, 245 (2011).
2. A. E. Renton *et al.*, *Neuron* **72**, 257 (2011).
3. N. J. Kramer *et al.*, *Science* **353**, 708 (2016).
4. R. Batra, K. Charizanis, M. S. Swanson, *Hum. Mol. Genet.* **19**, R77 (2010).
5. T. Zuo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 260 (2011).
6. C. R. Liu *et al.*, *Cell* **148**, 690 (2012).
7. H.-M. Cheng *et al.*, *PLOS Genet.* **11**, e1005043 (2015).
8. M. Wojciechowska, W. J. Krzyzosiak, *Hum. Mol. Genet.* **20**, 3811 (2011).
9. S. Mizielińska *et al.*, *Science* **345**, 1192 (2014).

10.1126/science.aah5408

Root cause

The *SUPT4H1*-*SUPT5H* complex binds RNA polymerase II (Pol II) and regulates transcription elongation of expanded nucleotide repeats, which cause a range of diseases including ALS. Targeting *SUPT4H1* reduces production of multiple toxic species, specifically sense and antisense repeat RNAs and repeat proteins.



SLEEP

Measuring the passage of brain time

Many brain areas are under both circadian and homeostatic control

By Charles A. Czeisler^{1,2}

The brain's central circadian (near-24-hour) clock exerts a very powerful influence on sleep and wakefulness (1, 2), interacting with our increasing drive to sleep as time awake increases (3, 4). On page 687 of this issue, Muto *et al.* (5) demonstrate that the interaction of the circadian clock with this so-called homeostatic sleep drive—which had been inferred from studies of human performance, sleep propensity, and risk of error or accident (3, 4, 6)—is rooted in measurable changes in cortical and subcortical responsiveness. Moreover, the nature of the interaction varies among cortical and subcortical brain regions.

Muto *et al.* analyzed the brain activity of 33 healthy individuals who were subjected to sleep deprivation for 42 hours under constant environmental conditions. During this time, brain activity was assessed every 2 to 6 hours (during particular tasks) by both electroencephalography and by functional magnetic resonance imaging (fMRI). The circadian component of the fMRI data was modeled by two different techniques. In one approach, sine and cosine waves (corresponding to the phases of the near-24-hour circadian clock) were fit to the fMRI data with standard techniques. This revealed significant circadian rhythmicity of brain responses in all brain regions except the dorsolateral prefrontal cortex, but with a peak in responsiveness that was a few hours earlier than anticipated from physiological and behavioral data (5). Fewer brain regions (subcortical regions such as the thalamus, head of the caudate nucleus, and putamen) showed significant circadian rhythmicity with a new but unconventional method that used the observed profile of melatonin production to fit the circadian component

¹Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, and the Sleep Health Institute, Brigham and Women's Hospital, Boston, MA 02115, USA.

²Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA. Email: charles_czeisler@hms.harvard.edu



Brain imaging reveals that both circadian (clock) and homeostatic (hourglass) control mechanisms interact to affect regional brain responsiveness over time.

of the fMRI data. Melatonin is secreted by the pineal gland in the brain and, when measured in constant dim light, reflects the output of the circadian clock. This new modeling method revealed that the phase of peak brain responsiveness is consistent with the timing of the evening “wake-maintenance” zone (a surge of wakefulness driven by the circadian clock that peaks just before the onset of melatonin secretion) (3, 7, 8). The authors’ conclusion that multiple brain areas with visually prominent daily variations in brain responsiveness lacked significant circadian rhythmicity as assessed by this unconventional technique is likely premature, given the limited statistical power available from the small number of individuals sampled. An alternative approach would have been to fit the fMRI data with more than one harmonic function to improve accuracy in estimating the phase of peak responsiveness, as has been done with melatonin data (9). However, the effort to fit a waveform was likely also hampered by the uneven intervals between fMRI assessments. Thus, it is likely that many more brain areas may yet be classified as being under both circadian (clock) and homeostatic (hourglass) regulation.

Although Muto *et al.* emphasized small differences in the timing of peak responsiveness from different brain regions, as estimated from relatively infrequent sample collection times, the similarity in the timing of peak responsiveness across regions is striking. The timing of all ~50 estimated circadian peaks measured in different brain regions occurred during only 17% of the possible range of circadian phases, validating conclusions derived from spontaneous desynchrony, forced desynchrony, and sleep deprivation protocols, for the existence of a wake-maintenance zone in the early evening (3, 7, 8). The peak in wake propensity presumably emanating from the suprachiasmatic nucleus (SCN) at that time not only opposes the increasing

drive for sleep that occurs near the end of the usual waking day, but is also timed to occur just before dusk, when most day-active species experience a surge in alertness that likely conferred an evolutionary advantage to diurnal animals seeking shelter from predators during their nighttime sleep. Although evening exposure to artificial light induces a substantial circadian phase delay in the timing of the wake-maintenance zone in most people living in industrialized countries (10), Muto *et al.* reveal that the peak of brain responsiveness is still timed to occur just before melatonin onset, which is a marker for the onset of biological dusk in humans. Similarly, the circadian nadir of brain responsiveness revealed in their study corresponds with a critical circadian zone of performance vulnerability in the early morning hours, previously demonstrated in controlled laboratory studies (4) and observed in motor vehicle crash statistics (6).

The SCN of the hypothalamus is the source of the circadian rhythm of responsiveness throughout the brain. However, even this new brain imaging study does not reveal whether the rise of sleep drive associated with increasing wakefulness is derived from a signal emanating from a specific location (a “sleep homeostat”), or instead represents a change in brain responsiveness that is due to the buildup of toxic metabolites (11), depletion of energy stores (12), or ionic changes in extracellular fluid (13) associated more broadly with sleep deficiency itself.

The findings of Muto *et al.* open many doors for addressing basic research questions. For example, the close temporal relationship between the circadian rhythm

of subcortical responses and the melatonin profile raises the question as to whether melatonin itself affects brain responsiveness, or whether it is simply serving as a marker of circadian phase. Assessment of the circadian rhythm of subcortical responsiveness in melatonin-deficient animals or humans may help to address this question. The experimental approach of Muto *et al.* also could be used to assess how the buildup of homeostatic sleep drive changes with time awake during brain development, particularly as the developing brain prunes or refines more than a hundred trillion synapses from pre- to postadolescence (14, 15). The approach could also reveal whether the decline in sleep capacity that occurs with advancing age reflects a decline in the need to sleep, or simply in the ability to do so. Perhaps the approach could even lead to the identification of a diagnostic biomarker that reflects homeostatic sleep drive. Once validated, such a technique could, for example, be used to distinguish between chronic sleep deficiency, which afflicts many children and adolescents, and attention deficit hyperactivity disorder, with which it shares many clinical symptoms.

Such a biomarker might also identify people suffering from chronic sleep deficiency whose fatigue is otherwise masked by caffeine and other stimulants. Overall, the study of Muto *et al.* opens up a new era in the quantitative assessment of sleep-wake neurobiology that links localized brain responsiveness to a

wide range of neurobehavioral performance and neuroendocrine and physiologic measures in humans, with applications in both health and disease. ■

“Overall, the study of Muto *et al.* opens up a new era in the quantitative assessment of sleep-wake neurobiology...”

REFERENCES

1. C. A. Czeisler *et al.*, *Science* **210**, 1264 (1980).
2. S. H. Strogatz, R. E. Kronauer, C. A. Czeisler, *Sleep* **9**, 353 (1986).
3. D.-J. Dijk, C. A. Czeisler, *Neurosci. Lett.* **166**, 63 (1994).
4. D. A. Cohen *et al.*, *Sci. Transl. Med.* **2**, 14ra13 (2010).
5. V. Muto *et al.*, *Science* **353**, 687 (2016).
6. J. F. Duffy, K. M. Zitting, C. A. Czeisler, *Rev. Hum. Factors Ergon.* **10**, 29 (2015).
7. S. H. Strogatz, R. E. Kronauer, C. A. Czeisler, *Am. J. Physiol.* **253**, R172 (1987).
8. J. A. Shekleton *et al.*, *J. Clin. Sleep Med.* **9**, 353 (2013).
9. E. N. Brown, Y. Choe, T. L. Shanahan, C. A. Czeisler, *Am. J. Physiol.* **272**, E506 (1997).
10. C. A. Czeisler, *Nature* **497**, S13 (2013).
11. L. Xie *et al.*, *Science* **342**, 373 (2013).
12. R. Basheer, R. E. Strecker, M. M. Thakkar, R. W. McCarley, *Prog. Neurobiol.* **73**, 379 (2004).
13. F. Ding *et al.*, *Science* **352**, 550 (2016).
14. I. Feinberg, H. C. Thode Jr., H. T. Chugani, J. D. March, *J. Theor. Biol.* **142**, 149 (1990).
15. E. P. Hoel, L. Albantakis, C. Cirelli, G. Tononi, *J. Neurophysiol.* **115**, 2199 (2016).

10.1126/science.aah5402

ATMOSPHERIC CHEMISTRY

Atmospheric radical chemistry revisited

Sunlight may directly drive previously unknown organic reactions at environmental surfaces

By Veronica Vaida

Sunlight is the largest energy source for Earth and therefore determines many aspects of our planet's chemistry and climate. For example, light-driven splitting (photolysis) of ozone at high altitude leads to the formation of hydroxyl radicals, which are involved in most oxidative processes in the environment. On page 699 of this issue, Rossignol *et al.* (1) report on an alternative process. They show that direct photolysis of a fatty acid at an air-water interface leads to the formation of oxidized products in the gas phase and of macromolecular products in water. This example, along with recently reported indirect photolysis of organic molecules (2, 3), shows

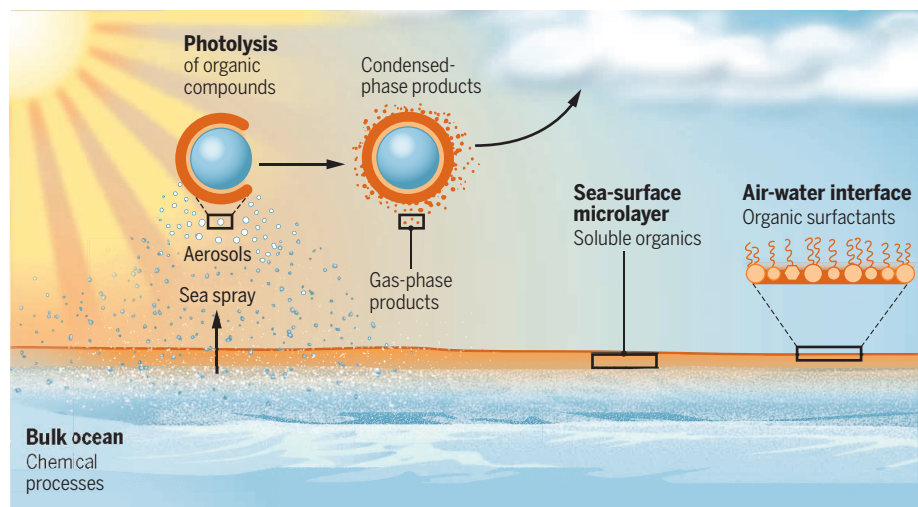
for Rossignol *et al.*'s observation of direct photolysis of nonanoic acid (1) is that the reaction occurs not in the gas phase but at a water-air interface. Such interfaces are found at the surfaces of oceans, lakes, cloud and fog droplets, and atmospheric aerosol particles (see the figure) (4–6). Hydrophobic organic molecules, such as the fatty acids and alcohols that are abundant in sea spray (7), concentrate at these interfaces (5, 6, 8, 9), where their properties are modified compared with the gas phase (4, 5, 10). Rossignol *et al.* show that when nonanoic acid is present in coatings at a water-air interface, it weakly absorbs ultraviolet light at wavelengths present in sunlight at Earth's surface. Gas- and condensed-phase organic photoproducts are observed as a result.

ate gas-phase and aqueous-phase functionalized complex organic species. The latter can form secondary organic aerosols in Earth's atmosphere (12). Organic radical recombination reactions can lead to the formation of oligomers, although competitive reactions involving oxygen addition may occur. At the water surface, however, where organic material is found in high concentration, radical initiation and recombination is expected to occur effectively, resulting in oligomers, polymers, and aggregates that in turn affect aerosol properties with consequences to climate.

Complex organic molecules and aggregates are known to form at the sea surface from biological processes (7). Recent findings (2, 3, 10–12), including those reported by Rossignol *et al.* (1), show that sunlight-initiated photochemistry contributes complex non-biological molecules at the sea surface (1, 3, 10, 11). These findings suggest that sunlight-initiated photochemical reactions at water surfaces have important consequences in the natural environment. However, questions remain about the generality of direct photolysis of organic compounds in the natural environment, calling for further studies of organic reaction mechanisms. The yields of direct and indirect organic photochemistry and the factors affecting their magnitude must be determined quantitatively before these reactions can be included in chemical models of the atmosphere. Nevertheless, the information now available suggests that sunlight-driven organic chemistry at the surface of water can produce high-molecular-weight products and aggregates. These products will affect secondary organic aerosol mass, composition, and optical properties, in turn defining the particle's overall effect on climate, air quality, and health. ■

Oxidation with a difference

Rossignol *et al.* report evidence for direct light-driven oxidation of an organic acid. Such oxidation processes may occur in the natural environment at water surfaces that are reached by solar radiation, generating gas-phase functionalized molecules and macromolecular condensed-phase products. These products affect aerosol formation and properties, influencing climate, air quality, and health.



that radical reactions initiated by absorption of sunlight can follow mechanisms previously unknown in Earth's atmosphere.

Current models assume that carboxylic acids and fatty alcohols are not directly affected by sunlight in Earth's atmosphere and that they are instead processed by reaction with hydroxyl radicals. The reason

Organic photochemical reactions of environmental interest in water are expected to proceed through the triplet state either by direct excitation, as is the case in nonanoic acid (1), or indirectly by either intersystem crossing from a strongly absorbing singlet state (2, 10) or energy transfer from a sensitizer (3, 11). After excitation of the triplet state, reactive organic radicals are produced. Not unlike the hydroxyl radical, these organic radicals then react with stable organic molecules to gener-

REFERENCES

1. S. Rossignol *et al.*, *Science* **353**, 699 (2016).
2. A. E. Reed Harris *et al.*, *J. Phys. Chem. A* **118**, 8505 (2014).
3. H. Fuet *et al.*, *J. Am. Chem. Soc.* **137**, 8348 (2015).
4. D. J. Donaldson, V. Vaida, *Chem. Rev.* **106**, 1445 (2006).
5. C. George, M. Ammann, B. D'Anna, D. J. Donaldson, S. A. Nizkorodov, *Chem. Rev.* **115**, 4218 (2015).
6. E. C. Griffith, A. F. Tuck, V. Vaida, *Acc. Chem. Res.* **45**, 2106 (2012).
7. X. Wang *et al.*, *ACS Cent. Sci.* **1**, 124 (2015).
8. M. T. C. Martins-Costa, J. M. Anglada, J. S. Francisco, M. F. Ruiz-Lopez, *J. Am. Chem. Soc.* **134**, 11821 (2012).
9. R. Vacha, P. Slavicek, M. Mucha, B. J. Finlayson-Pitts, P. Jungwirth, *J. Phys. Chem. A* **108**, 11573 (2004).
10. E. C. Griffith, R. J. Rapf, R. K. Shoemaker, B. K. Carpenter, V. Vaida, *J. Am. Chem. Soc.* **136**, 3784 (2014).
11. R. Ciuraru *et al.*, *Sci. Rep.* **5**, 12741 (2015).
12. P. Renard *et al.*, *J. Phys. Chem. C* **118**, 29421 (2014).

10.1126/science.aah4111

Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309, USA. Email: vaida@colorado.edu



The Morley Nelson Snake River Birds of Prey National Conservation Area in Idaho preserves a wildlife habitat while also providing for compatible, recreational use of the land.

CONSERVATION

Balancing economic and ecological goals

What are the trade-offs between economic development and ecosystem conservation?

By Eyal G. Frank¹ and Wolfram Schlenker^{1,2}

Human activities are increasingly degrading ecosystems, resulting in habitat loss and reduced biodiversity. Extinction rates, a widely used metric of biodiversity loss, are estimated to be around 1000 times as high as those historically experienced on Earth (1); even the most conservative estimate puts current extinction rates at 114 times as high as the background planetary norm (2). Simulations project that, under scenarios of increased economic growth and the accompanying land use change, critical habitat will further degrade and biodiversity will decline (3, 4). Are economic growth and ecosystem conservation incompatible objectives?

HOW TO VALUE ECOSYSTEMS

In theory, economic growth is not at odds with preserving ecosystems (5). Economic analysis considers two classes of values for an ecosystem: Use values capture the direct benefits from ecosystem functions such as pest control, whereas existence values reflect the desire to preserve species even if they provide no direct function to the person who values them. If markets correctly price all inputs, economic incentives will

encourage preservation—for example, by becoming more efficient in the use of inputs that are costly.

But in reality, there are two daunting challenges. First, it is empirically very difficult to estimate existence values, and there is an active debate about whether they can be reliably estimated at all (6). Yet, refusing to incorporate any existence values because of the empirical challenges in estimating them in effect assumes that their value is zero.

Second, use values might not be priced correctly because they often benefit parties beyond the ones who provide and pay for them. The benefits to others need to be reflected in prices or the party providing biodiversity conservation will protect too little because it does not get rewarded for those benefits. For example, “Earth Sanctuaries,” a publicly listed for-profit company cited in (5) that invested money in preserving ecosystems with an expectation of making profits, no longer exists. The likely reason is that governments did not institute the required price signals and restrictions to make private conservation profitable. Governments must either set the correct price signals or directly place areas under protection.

PUTTING AREAS UNDER PROTECTION

Instituting the optimal policy is particularly challenging in developing countries. They often have weak institutions for enforcing regulation and are more susceptible to individuals exploiting their political

connections (7). They might also find it unappealing to impose penalties on poor constituents. Instead, both governmental and nongovernmental agencies have opted to pay for ecosystem services, but recent evidence suggests that these payments work best when poverty is low (8). Farmers in immediate need of money and without access to credit markets might choose to deforest because the one-time, immediate value of the harvest trumps larger combined future payments (9). Payments for ecosystem services have been shown to slow deforestation but did not reverse or halt it (10). These findings should be taken into account in the implementation of the Green Climate Fund, which is supposed to provide 100 billion dollars annually to fund low-emission and climate-resilient development such as avoided deforestation.

Protected areas present a trade-off between the lost economic activity and the benefits of conservation. Comparison of economic performance and protection levels around the world shows that, overall, richer countries have a higher share of their area under protection (see the figure). After the Earth Summit of 1992, interest in biodiversity grew dramatically (11). Countries in Europe and North America saw a sharp rise in protected area from 1992 to 2012. For the same time period, the area protected in Asian, Latin, and South American countries increased little despite remarkable growth in gross domestic product (GDP) per capita.

¹Columbia University, New York, NY, USA. ²National Bureau of Economic Research, Cambridge, MA, USA. Email: eyal.frank@columbia.edu; wolfram.schlenker@columbia.edu

Growth in both GDP and area protected was limited in African countries, although some of them had remarkably high shares of protected area to start with in 1992. Thus, there is no one-to-one mapping between economic growth and land protection.

BALANCING PROTECTION AND GROWTH

Knowing both how to balance protection relative to economic growth and how individuals and firms respond to additional protection is of key interest when designing policies. Economists have long argued that secure properties are crucial so that people face the right incentive to conserve. In fisheries, implementation of individual transferable quotas has corrected private incentives and slowed species extinction (12). An alternative for nonmigratory species is to create exclusive fishing zones that give the right incentive to manage and protect fish species.

The spatial extent of the protected area is also important; contiguous protected areas are crucial for preserving species that require a certain habitat size for survival. Yet, the standard approach so far (8–10) has been to pay for each protected acre irrespective of how the preserved area fits into the larger landscape. Because of such spatial interdependencies, optimal prices and payments need to vary—for example, by taxing pollution differently depending on where it is emitted (13). The problem is amplified in ecosystem conservation under climate change, which alters the niches that species occupy (14). Effective conservation under climate change may require contiguous corridors, along which regulatory approaches differ.

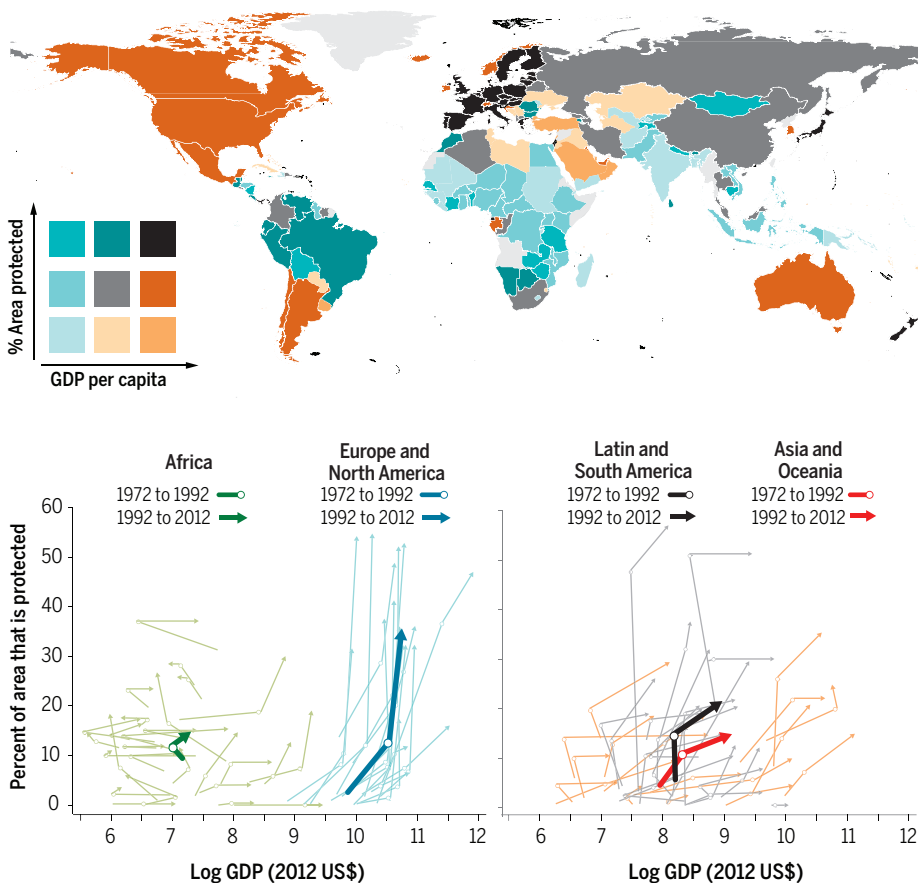
The optimal policy must also be mindful of unintended consequences. Regulation must be comprehensive to avoid “leakage,” where protected areas are offset by increased destruction in other places. The

sequence of events and the time window for regulated entities to react are crucial. For example, the U.S. Endangered Species Act requires several steps to get a species protected. This creates strong incentives for a landowner to eliminate the species before it becomes protected, a possible reason that adding a species does not decrease the probability of extinction (15).

Reductions in ecosystem functions can to some extent be compensated with human-made capital. However, these counteracting adjustments might negatively affect others. For example, increased use of fertilizer can compensate for soil degradation but can damage aquatic life and have negative health effects. Given that substitutability is a key concept in the economic perspective on sustainable development, there is a critical need for more empirical studies that use observational data and credible statistical analysis to estimate the effects of compensating a decline in ecosystem functions with human-made capital.

THE INFLUENCE OF GROWTH AND EQUALITY

Economic growth is a crucial ingredient for preservation but does not guarantee it. As people become richer and are less preoccupied with their survival, the values placed on preservation tend to increase, and richer societies will likely improve enforcement. By the same token, if preservation values increase with income but at a decreasing rate, as commonly assumed, then a more equal society will exhibit higher values for conservation. The income distribution might thus be as important as overall economic growth. ■



Wealth and conservation by country. (Top) This map compares GDP per capita in 2012 to the share of a country's area that was protected in the same year. High-income countries—e.g. those in North America and Europe—generally have higher levels of protection. However, some low-income regions also have high levels of protection, particularly in South America and Africa. We do not analyze how effectively land protection is implemented in various countries and do not claim any causality. **(Bottom left and bottom right)** The relationship between per capita GDP and protected area changed in many countries around the world between 1972 and 2012. We divide the data into 1972 to 1992 and 1992 to 2012. The richest countries have seen the largest increase in protected area after 1992, the year of the Earth Summit in Rio de Janeiro. See the supplementary materials for information on data sources.

REFERENCES

1. S. L. Pimm et al., *Science* **344**, 987 (2014).
2. G. Ceballos et al., *Sci. Adv.* **1**, e1400253 (2015).
3. T. Newbold et al., *Nature* **520**, 45 (2015).
4. E. F. Lambin, P. Meyfroidt, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3465 (2011).
5. G. C. Daily et al., *Science* **289**, 395 (2000).
6. C. L. Kling, D. J. Phaneuf, J. Zhao, *J. Econ. Perspect.* **26**, 3 (2012).
7. R. Burgess, M. Hansen, B. A. Olken, P. Potapov, S. Sieber, *Quart. J. Econ.* **127**, 1707 (2012).
8. J. M. Alix-Garcia, K. R. E. Sims, P. Yañez-Pagans, *Am. Econ. J. Econ. Policy* **7**, 1 (2015).
9. S. Jayachandran, *Am. Econ. Rev. Pap. Proc.* **103**, 309 (2013).
10. S. Jayachandran, J. de Laat, E. F. Lambin, C. Y. Stanton, *NBER Working Paper* 22378, June 2016.
11. B. J. Cardinale et al., *Nature* **486**, 59 (2012).
12. C. Costello, S. D. Gaines, J. Lynham, *Science* **321**, 1678 (2008).
13. N. Z. Muller, R. Mendelsohn, *Am. Econ. Rev.* **99**, 1714 (2009).
14. S. Lavergne, N. Mouquet, W. Thuiller, O. Ronce, *Annu. Rev. Ecol. Evol. Syst.* **41**, 321 (2010).
15. P. J. Ferraro, C. McIntosh, M. Ospina, *J. Environ. Econ. Manage.* **54**, 245 (2007).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/651/suppl/DC1

10.1126/science.aaf9697



GLOBAL CHANGE RESEARCH

The challenge of climate-change neoskepticism

Decision science and risk management are underutilized

By Paul C. Stern,^{1,2*} John H. Perkins,^{3,4}
Richard E. Sparks,⁵ Robert A. Knox⁶

Opponents of policies to limit anthropogenic climate change (ACC) have offered a changing set of arguments—denying or questioning ACC's existence, magnitude, and rate of progress, the risks it presents, the integrity of climate scientists, and the value of mitigation efforts (1). Similar arguments have characterized environmental risk debates concerning arsenical insecticides in the late 1800s (2), phosphates in detergents in the 1960s (3), and the pesticide DDT in the 1960s and '70s (4). Typically, defenders of business as usual first question the scientific evidence that risks exist; then, they question the magnitude of the risks and assert that reducing them has more costs than benefits. A parallel rhetorical shift away from outright skepticism (5–7) led us to identify “neoskepticism” as a new incarnation of opposition to major efforts to limit ACC (8). This shift heightens the need for science to inform decision making under uncertainty and to improve communication and education.

In the historical cases above, dispute resolution rested on both finding new, less-risky practices and applying policy judgment to science. In the DDT case, the administrator of the U.S. Environmental Protection Agency (EPA) stated that he was “convinced by a preponderance of the evidence” (4). Today, neoskepticism accepts the existence of ACC but advocates against urgent mitigation efforts on various grounds, such as that climate models run “too hot” (5) or are too uncertain to justify anything other than “no-regrets” policies as having net benefits (6). Mainstream climate scientists are well aware of uncertainty in climate projections (9, 10). But neoskeptics’ citing of it to justify policy inaction marks a shift of focus in climate debates from the existence of ACC to its import and to response options.

LEGITIMATE QUESTIONS, UNJUSTIFIED INFERENCES

Neoskeptical claims may be driven more by ideology or economic interests than by science (11), but they cannot be dismissed as confidently as pure denial or skepticism. By focusing attention on risks and highlighting uncertainty, they raise important questions about ultimate impacts and response options. These are questions about (i) the nature and probabilities of crossing climate “tipping points” that would greatly change the profile of the risks (12); (ii) the varied harms climate change may produce (e.g., to businesses, ecosystem integrity, political

stability, or human lives); (iii) the vulnerability and resilience of potentially affected social and ecological systems; and (iv) the benefits and costs of efforts to limit climate change, reduce vulnerability, and increase resilience. Such questions have received scientific and economic analysis before (9) but deserve greater emphasis as attention focuses increasingly on risk management.

Although neoskeptics claim to accept the reality of ACC, their inference that inaction is justified seriously underemphasizes some well-established characteristics of ACC that are important for informing choices: that the risks of extreme and damaging outcomes are continually increasing, so that waiting for certainty has increasing costs; that inertia in the system may result in its crossing major tipping points without timely warning; and that there is value to insuring against worst cases, especially when they are likely to be worse than those of the past. Only by presuming that the risk profile is static can neoskeptics plausibly argue that uncertainty justifies postponing action and that mitigation efforts that do not pay off well now will be regretted later. A more appropriate inference is that efforts that look overly costly now may well appear in retrospect to have been low-cost insurance.

SCIENCE FOR DECISION MAKING

Science for climate decisions under uncertainty requires shifts in research agendas. For informed responses to ACC, science must not only assess climate trends and project their consequences for diverse affected parties but also (i) identify specific technological and institutional options; (ii) analyze the time and resources required for mitigation and adaptation strategies; (iii) analyze the effects of each and of ways to increase effectiveness; (iv) identify and assess cobenefits and indirect costs of mitigation; and (v) evaluate the financial costs and other risks of delay. The science of decision processes is also needed. Much of the needed research must come from the social, behavioral, and economic sciences integrated into a science of human-environment interactions (9, 13).

By acknowledging ACC but questioning the value of mitigation efforts, neoskepticism highlights the need for science to inform adaptive risk management (9) and for serious risk analysis. Deterministic analysis has limited potential for comparing the costs of actions versus delay because the probabilities of occurrence of the most damaging conditions, although changing and increasing with delay, will long remain uncertain, and these probabilities are the main physical variable affecting climate risks. The science must consider known and potential risks and uncer-

¹National Research Council, Washington, DC 20001, USA.

²Norwegian University of Science and Technology, 7491

Trondheim, Norway. ³The Evergreen State College, Olympia, WA

94708, USA. ⁴University of California, Berkeley, CA 94720, USA.

⁵University of Illinois, Champaign, IL 61801, USA. ⁶University of

California, San Diego, La Jolla, CA 92093, USA. *Email: pstern@nas.edu

tainties and must assess options for reducing risks and for altering plans and strategies as new information becomes available.

The science needed to integrate information about decision options and climate risks has not been well represented in the Intergovernmental Panel on Climate Change (IPCC) process (14) or adequately supported in research budgets. For example, despite continued calls from independent reviews for increased support for “human dimensions” research in the U.S. Global Change Research Program, funds devoted to social science declined from 3 to 2% between 1990 and 2007 (15). The program’s most recent strategic plan embraces science to inform timely decisions on adaptation and mitigation, but resources commensurate with the needs have not been achieved.

Decision science is particularly needed, both because of uncertainties about risks, costs, and benefits and because of differing information needs and value judgments within societies about the relative importance of different outcomes, affected populations, and time horizons. Catastrophic outcomes are the most uncertain and will likely elicit conflicts about priorities. The decision sciences offer useful guidance for informing decisions under these conditions. With input from decision participants, these sciences can assess—in ways relevant to citizens—the costs, feasibility, and side effects of the response options, including inaction. Decision sciences can also suggest ways of organizing decision making to integrate scientific analysis and social deliberation and to consider trade-offs that might otherwise go unaddressed. In addition, they can incorporate issues of equity and value conflict in supporting adaptive risk management (16).

The decision sciences suggest broad principles for uncertainty management, such as (i) adopting policies that will perform robustly across various plausible futures, (ii) pursuing a variety of policy strategies to increase the likelihood that some will yield good results, and (iii) organizing decision-making processes for flexibility and responsiveness (16). Monitoring and assessment of results are essential for identifying policy improvements for adaptive management. All these activities of decision science help the public and decision makers assess the options for action and the consequences of inaction—the main challenge unaddressed by neoskepticism.

IMPLICATIONS FOR SCIENCE COMMUNICATION AND EDUCATION

Neoskeptical arguments implicate but do not appropriately address the cognitively difficult issues of understanding scientific uncertainty and weighing value trade-offs. To cope with such difficulties, most people use

simple mental models (17). They need simple models that incorporate the key elements of the current mixture of established knowledge and uncertainty, can suggest decision strategies, and can adapt to emerging knowledge and experience. Skeptics and neoskeptics, however, often suggest simple models that are misleading because they do not take into account the continually increasing costs of waiting for certainty.

Neoskepticism thus challenges climate educators and communicators to supplement the teaching of climate facts and processes with mental models of ACC that (i) are factual and not misleading, (ii) use a familiar domain to explain the unfamiliar, (iii) capture interest, and (iv) allow for extrapolations consistent with current science. Because risks and value trade-offs are key, simple mental models should also (v) acknowledge and incorporate uncertainty, (vi) encourage choices about risks in the face of uncertainty, (vii) highlight unresolved issues for discussion among people who may initially disagree,

“Neoskeptical arguments ... do not appropriately address ... scientific uncertainty and weighing value trade-offs.”

and (viii) highlight the accumulating cost of waiting for certainty (18).

Citizens and nonscientist policy-makers are quite capable of considering serious hazards whose probabilities are uncertain, as they do with hazards as diverse as fires, floods, extreme storms, airline crashes, infectious disease outbreaks, terrorist attacks, and impending wars. Many of these risks are managed to reasonable levels of public satisfaction even though neither their probabilities of occurrence nor the benefit-cost ratios of risk management options can be estimated with precision and confidence.

Some familiar decision-making settings may capture fundamental attributes of climate change and serve as helpful analogs. For example, coping with risks of progressive medical conditions, such as hypertension or atherosclerosis, involves choosing whether to mitigate risks through changes in diet, smoking, and exercise without knowing how much change is needed. Like climate change, these medical conditions are progressive, can lead to catastrophic tipping points (stroke or heart attack), and are only slowly reversible. Decisions about the purchase of insurance against catastrophic life events (illness, fires, floods, and others) also offer a useful analog.

Each analogy will suggest choice options. Although there is no perfect familiar analog for climate change, we suggest that some simple mental models, by making key fundamentals of climate change more understandable, can facilitate thinking about the choices facing society. They may also have potential to improve public discourse addressing risks of climate change by making clear that certain kinds of policy disagreements should be expected, even with improved knowledge.

CONCLUSION

We do not presume that empirical analysis of risks, techniques for informing decisions under uncertainty, or better mental models will end skepticism, neo- or otherwise, about climate change. Some skepticism is motivated by ideological and financial interests tied to the continued use of fossil fuels, and debates will continue over the risks and how best to manage them. Nevertheless, improved understanding of the issues raised here can assist decision making in the face of the risks of climate change and can perhaps help focus policy debates more productively. ■

REFERENCES AND NOTES

1. C. Boussalis, T.G. Coan, *Glob. Environ. Change* **36**, 89 (2016).
2. J. Whorton, *Before Silent Spring: Pesticides and Public Health in Pre-DDT America* (Princeton Univ. Press, Princeton, NJ, 1974), pp. 20–35.
3. J.R. Vallentyne, “The algal bowl: Lakes and man” (Misc. Special Publ. 22, Department of the Environment, Fisheries and Marine Service, Ottawa, 1974).
4. EPA, DDT: A Review of Scientific and Economic Aspects of the Decision to Ban Its Use as a Pesticide, prepared for Committee on Appropriations, U.S. House of Representatives (EPA-540/1-75-022, EPA, Washington, DC, 1975).
5. C. Monckton et al., *Sci. Bull.* **60**, 122 (2015).
6. S. Koonin, *Wall Street J.*, 19 September 2014.
7. J. Curry, *Wall Street J.*, 9 October 2014.
8. J.H. Perkins et al., *Nature* **522**, 287 (2015).
9. National Research Council (NRC), *Advancing the Science of Climate Change* (National Academies Press, Washington, DC, 2011).
10. IPCC, *Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, et al., Eds. (Cambridge Univ. Press, Cambridge, 2013).
11. J. Farrell, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 92 (2016).
12. NRC, *Abrupt Impacts of Climate Change: Anticipating Surprises* (National Academies Press, Washington, DC, 2013).
13. P.C. Stern et al., *Nat. Clim. Change* **10**, 1038/NCLIMATE3027 (2016).
14. C. Carraro et al., *Science* **350**, 34 (2015).
15. NRC, *A Review of the U.S. Global Change Research Program’s Draft Strategic Plan* (National Academies Press, Washington, DC, 2012).
16. NRC, *Informing Decisions in a Changing Climate* (National Academies Press, Washington, DC, 2009).
17. N.F. Pidgeon, B. Fischhoff, *Nat. Clim. Change* **1**, 35 (2011).
18. P.C. Stern, K.T. Raimi, *Soc. Res. Int. Q.* **82**, 583 (2015).

ACKNOWLEDGMENTS

Except as noted in references, the views expressed in this paper are those of the authors and not of the National Academies of Sciences, Engineering, and Medicine. J.H.P. thanks S. Jenkins and C. Merchant for sponsoring him as a Visiting Scholar at University of California, Berkeley. We also thank T. Dietz and two anonymous reviewers for comments.

10.1126/science.aaf6675



A potent combination of anticoagulants and vasodilators amplifies the lethality of the Komodo dragon's bite.

BIOCHEMISTRY

Dangerous liaisons

An insider's guide to venomous creatures delivers on the science but is short on tales of the bitten and stung

By Christopher Kemp

In the 2004 Discovery Channel documentary “Killer Jellyfish,” Australian toxinologist Jamie Seymour is stung by an Irukandji jellyfish. A box jellyfish, the Irukandji is pea-sized and transparent. It drifts through the water like a ghostly mushroom. For hours after he is stung, Seymour writhes on a hospital bed, his chest covered with EKG leads. His legs tremble. His agony lasts for 50 hours. In others it lasts longer—2 weeks or more. And, sometimes, it's fatal.

An evolutionary biologist at the University of Hawaii, Christie Wilcox studies the evolution of toxins in lionfish, scorpionfish, and species like the Irukandji jellyfish. With *Venomous*, she takes us on a biochemical tour of the natural world, examining how certain species have evolved the ability to defend and kill with venom.

The honey bee uses a toxin called melittin. Both the blue-ringed octopus and the pufferfish use tetrodotoxin—a potent neurotoxin that inhibits the firing of action potentials in nerves. The male duck-billed

platypus unleashes venom from 1-inch-long toxic spurs on its ankles; its production requires 83 different genes. Covered in bristles, the *Lonomia* caterpillar delivers a substance that begins with hemorrhagic shock and ends with organ failure. Vipers use hemotoxins: the body bleeds to death inside. The cone snail—*Conus geographus*—doesn't sting many people, but it kills 70% of those it does. Generally, bites are offensive and stings are defensive.

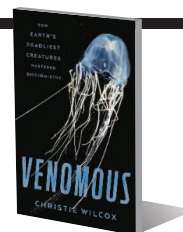
Despite all of our best efforts, writes Wilcox, the effects of bites and stings are still a problem worldwide. In India alone, the four most venomous snake species—the Russell's viper, the saw-scaled viper, the spectacled cobra, and the common krait—still manage to kill tens of thousands of people every year. Usually, they kill poor people in remote places. The antivenom exists, but it is inaccessible.

Interestingly, the diversity of venomous strategies across species is not as great as one might at first think. Some very different species have evolved identical toxins, acting on the same internal systems and cellular pathways. Snakes, jellyfish, and spiders all produce tissue-destroying phospholipase A2. Snails, corals, and worms all produce Kunitz-type peptides, which prevent other proteins from functioning properly. Some

Venomous How Earth's Deadliest Creatures Mastered Biochemistry

Christie Wilcox

Scientific American/Farrar, Straus, and Giroux, 2016. 251 pp.



species, particularly the cone snails, recruit a cocktail of hundreds of venomous compounds, but, even then, many of them are the usual suspects.

From an evolutionary point of view, if an idea is good enough, more than one species will have it, and some species will have the same idea again and again. Reptiles have recruited phospholipase A2 at least four different times in their evolutionary history.

As Wilcox explains all this, she travels to Tasmania to observe a duck-billed platypus; treks into the Peruvian jungle to find the dreaded *Lonomia* caterpillar; and flies to remote Indonesian islands to observe the Komodo dragon, whose venom acts as a potent anticoagulant and causes a rapid decrease in blood pressure that induces shock. We meet toxinologists who study venoms and also self-immunizers—a small group of laypeople who willingly inject themselves with diluted snake venom, reporting that cobra venom gives them a high that is better than cocaine.

The subject matter Wilcox is writing about sits at a fascinating intersection between the human and the natural worlds. We are interested in these deadly animals—the jellyfish in the water; the camouflaged scorpion beneath the rock, poised to strike—precisely because of what they might do to us. The idea of being bitten by a rattlesnake in the desert, miles from anywhere, fills us with dread.

Wilcox tells us that thousands of people are bitten and stung each year. So, where, I wondered, are the testimonies of the snakebitten? Where is the desperate and dry-throated rush across state lines for a vial of antivenom as the bitten hand turns into a shiny red balloon? For a subject with such a human impact, I was constantly wondering: Where are all the people?

Our understanding of the biochemistry of venoms, and how they evolved, is still incomplete. The field of venomics, the study of naturally occurring venoms as a novel drug discovery platform, is in its infancy, too. Wilcox does a good enough job of introducing us to them both. But the dread—the leg-trembling agonies Jamie Seymour endured after his Irukandji sting—must be looked for elsewhere.

10.1126/science.aag2368

MEDIA STUDIES

In real life

An essayist offers an intimate glimpse into the digital technologies that are redefining the human experience

By Jennifer Golbeck

Laurence Scott was born in the twilight zone between generation Xers and millennials, coming of age as social media emerged on the web and changed the world. *The Four-Dimensional Human* is part tech-philosophical treatise and part literary memoir, exploring how constant connectivity has changed us and our interactions.

Scott's 4D human lives in the physical world and in the ethereal online space, occupying both at the same time. Online interactions are free of physical constraints, and time operates differently there, but it inevitably comes back to affect our offline lives. The strangeness and tension that emerge from this is the core theme of the book.

Life in the fourth dimension is most eloquently captured in the book's opening example, in which Scott describes being out for the evening in Dublin with his college classmates and tutor. They descend into a basement wine bar, full of books, flowers, art, and gilt. The tutor orders champagne, which arrives at the same time Scott receives a text from a new connection. He busily responds, missing the toast. This state of being physically in one place but drawn, through technology, to somewhere else is an eminently relatable experience.

Scott does not advance an argument so much as he shares reflections on a bouquet of topics, moving from one to another like a bee moves between flowers. They include some that you might expect: how technology has enabled changes in language, the emergence of "sharing" culture, the appropriation of ideas, and new ways of dealing with boredom.

But he also ranges more broadly, comparing the emergence of colony collapse disorder in bee populations with our "online migration ... to a virtual hive," arguing that Gap's 2014 "Dress Normal" campaign tapped into our collective anxiety about modes of self-expression in the digital age, and chronicling a failed attempt to replace time zones with a decimalized system of a thousand beats known as "Internet

Time." His conclusions tend to be that our "fourth-dimensional" experiences are often strange, anemic, or disconcerting, but he leaves hope that we can find a way of living happily in this new world.

Scott has clearly thought deeply about all the issues he explores. The book is carefully crafted and eloquently written. It is also a refreshing departure from the libraries of tech books written by self-proclaimed futurists filled with bold claims and (empty) promises about how technology is changing the world.



"... [I]f you dabble in other realities, then you shouldn't expect to remain unchanged," writes Laurence Scott.

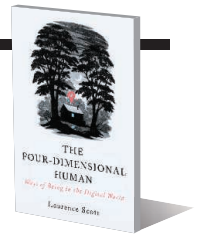
That said, the book truly is part memoir, and some of Scott's experiences are far from universal. He spends a few pages early on describing "the strange sorrow that Skype provokes," a sense of isolation that he finds distinct and profound when compared with phone calls. Although it may be an experience shared by some, this is certainly not a universal reaction. Research has shown that for many people, Skype actually reduces the feeling of distance (1–3).

He also writes at length about Airbnb, describing it as voyeuristic for the way it lets us see into others' homes and also as evoking the uncanny valley for giving us a feeling of home when we are away. But people were renting houses, complete with listings

The Four-Dimensional Human

Ways of Being in the Digital World

Laurence Scott
Norton, 2016. 272 pp.



with pictures, well before the Internet, and I was left unconvinced that moving this industry online has produced any fundamental change in the experience.

Let's cut to the chase: Are you going to enjoy this book or not? The answer is a big "it depends"; some people will love it and others will have strong feelings in the other direction. Fortunately, you can quickly find out what camp you are in with this simple test.

Consider this passage that deals with originality online: "The digital, in one strike, manages to sublimate the real and petrify

the ethereal. As a result, we feel our lives evaporating and solidifying in the same breath, and as when something very hot meets something very cold, we feel cracks beginning to form."

Like it? Then you will love this book. But whether or not you add *The Four-Dimensional Human* to your summer reading stack, it is a valuable book that brings a fresh perspective to the topic of life online and offline.

REFERENCES

1. R. Longhurst, *Environ. Plan D*, **31**, 4 (2013).
2. R. C. King-O'Riain, in *Internet and Emotions* (Routledge, Abingdon, 2013).
3. A. Aguila, *Exp. Media Ecol.*, **10**, 3 (2012).

10.1126/science.aag1305

The reviewer is at the Social Intelligence Lab, University of Maryland, College Park, MD 20742, USA. Email: golbeck@cs.umd.edu

LETTERS

Edited by Jennifer Sills

China's ecosystems: Focus on biodiversity

IN THEIR REPORT "Improvements in ecosystem services from investments in natural capital" (17 June, p. 1455), Z. Ouyang *et al.* state that China's investment in natural capital has contributed substantially to improvements in most ecosystem services, with one exception: Habitat provision for biodiversity decreased slightly between 2000 and 2010. The implications of this slight decrease in biodiversity conservation should not be overlooked.

The latest Red Lists of China's biodiversity show that ~22% of vertebrates and ~11% of higher plants are either extinct or threatened (1, 2). The rate of biodiversity loss has not slowed (3). Habitat provision remains vital to conservation. Ouyang *et al.* explain that China's environmental investment programs "aim to reduce natural disaster risk by restoring forest and grassland," but this goal lacks a direct link to biodiversity conservation. For instance, nonnative planted forests can provide important ecosystem services but contribute very little to regional biodiversity.

To enhance biodiversity conservation, ecosystem service policies should incorporate goals specific to biodiversity; focus on implementation in areas that integrate both biodiversity and ecosystem services (4), as identified by systematic conservation planning; emphasize the role of existing protected areas in ecosystem services provision; and prioritize remaining intact ecosystems and wilderness areas. These areas are often the last refuges of endangered and endemic species. They also provide irreplaceable ecosystem services (5).

Ruidong Wu

Institute of International Rivers and Eco-Security and Yunnan Key Laboratory of International Rivers and Transboundary Eco-Security, Yunnan University, Kunming, Yunnan 650091, China.
Email: rdwu@ynu.edu.cn

REFERENCES

1. Ministry of Environmental Protection of China, Chinese Academy of Sciences, "Assessment Report on the Red List of China's Biodiversity—Higher Plants" (2013); www.mep.gov.cn/gkml/hbb/bgg/201309/t20130912_260061.htm [in Chinese].
2. Ministry of Environmental Protection of China, Chinese Academy of Sciences, "Assessment Report on the Red List of China's Biodiversity—Vertebrates" (2015); www.zhb.gov.cn/gkml/hbb/bgg/201505/t20150525_302233.htm [in Chinese].

Red-necked pond turtle (*Mauremys nigricans*)



3. S. H. M. Butchart *et al.*, *Science* **328**, 1164 (2010).
4. R. Wu *et al.*, *PLOS ONE* **9**, e103783 (2014).
5. S. Lin *et al.*, *Sci. Rep.* **6**, 25898 (2016).

10.1126/science.aah3994

China's ecosystems: Overlooked species

A RECENT REPORT by Z. Ouyang *et al.* ("Improvements in ecosystem services from investments in natural capital," 17 June, p. 1455) found that investment in the restoration and preservation of natural capital in China has brought improvements at the national level in most of the major ecosystem services measured. Our recent status assessment of China's vertebrates (1) tells a less hopeful story.

Although charismatic species, such as the giant panda (*Ailuropoda melanoleuca*) and crested ibis (*Nipponia nippon*) have been rescued from the brink of extinction (1), species that do not benefit from media attention are suffering. Seventeen vertebrate species are extinct or regionally extinct, including the Yunnan lake newt (*Hypselotriton wolterstorffi*) and the Caka stone loach (*Triplophysa cakaensis*). The red-necked pond turtle (*Mauremys nigricans*) and the Liaoning clawed salamander (*Onychodactylus zhaermii*) are among the 43% of China's vertebrates that are in peril. Some may disappear soon if urgent action is not taken (1). The proportion of threatened amphibians, also overlooked by the media, is as high as 63% (1). As indicated by the Red List Index (2), the status of mammals and fishes in China is deteriorating rapidly, amphibians and reptiles are threatened, and many birds lack sufficient data to determine a status.

Existing laws protect a few charismatic species (3), but all species need legal protection. Although most are not deemed worthy of media headlines, all species are

crucial to the stability and function of ecosystems (4).

Zhigang Jiang,^{1*} Jianping Jiang,² Yuezhao Wang,² E. Zhang,³ Yanyun Zhang,⁴ Lili Li,¹ Bo Cai,² Zhenhua Luo,⁵ Chunwang Li,¹ Xiaoge Ping,¹ Feng Xie,² Liang Cao³

¹Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. ²Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan 610041, China. ³Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei 430000, China. ⁴Life Science College, Beijing Normal University, Beijing 100875, China. ⁵School of Life Sciences, Central China Normal University, Wuhan, Hubei 430079, China.

*Corresponding author: Email: jiangzg@ioz.ac.cn

REFERENCES

1. Z. Jiang *et al.*, *Biodiv. Sci.* **24**, 552 (2016).
2. IUCN, "IUCN Redlist" (2016); www.iucn.redlist.com.
3. State Council of PR China, *National Key Protected Wild Animal Species* (2003); www.forestry.gov.cn/portal/bhxs/s/709/content-85157.html [in Chinese].
4. R. P. Leitão *et al.*, *Proc. R. Soc. Ser. B Biol. Sci.* **283**, 10.1098/rspb.2016.0084 (2016).

10.1126/science.aah4000

China's ecosystems: Sacrificing the poor

IN THEIR REPORT "Improvements in ecosystem services from investments in natural capital" (17 June, p. 1455), Z. Ouyang *et al.* model changes in selected ecosystem services in China to show that, at the national level, rapid economic growth is compatible with increased provision of ecosystem services. However, closer examination of their data reveals that national-scale improvements have come at the cost of environmental justice toward poor, marginalized, and ethnic minority communities, whose well-being is sacrificed to provide ecosystem services for wealthier, mostly urban populations.

For example, Sanjiangyuan, a grassland region found to have a net increase in ecosystem services, has been subject to

purported grassland restoration and ecological migration programs that destock and move Tibetan herders off the land to new and often poorly built settlements, where they face health problems, unemployment, declines in living standards, and loss of Tibetan language and cultural practices (1, 2). Their livelihoods and rights are sacrificed to provide downstream ("national") ecological security.

Furthermore, reforestation programs have increased forest cover at a national scale, but many plantations have failed because of inappropriate species selection or inadequate financial compensation to farmers recruited to plant trees. Others have yielded forests that provide downstream ecosystem services but bear few if any benefits locally (3). China also makes up for lost lumber production by importing timber, often illegally harvested in poorer countries, contributing to global-scale environmental injustice by exporting deforestation (4). Given that degradation and disasters harm human populations, efforts to steward resources must be equitable.

John A. Zinda,^{1*} Kelly A. Hopping,² Edwin Schmitt,³ Emily T. Yeh,⁴ Stevan Harrell,⁵ Eugene N. Anderson⁶

¹Department of Development Sociology, Cornell University, Ithaca, NY 14853, USA. ²Department of Earth System Science, Stanford University, Stanford, CA 94305, USA. ³Department of Anthropology, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China. ⁴Department of Geography, University of Colorado, Boulder, CO 80309, USA. ⁵Department of Anthropology and School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195, USA. ⁶Department of Anthropology, University of California, Riverside, CA 92521, USA.

*Corresponding author. Email: jaz65@cornell.edu

REFERENCES

1. F. Du, *Nomadic Peoples* **16**, 116 (2012).
2. J. M. Foggin, *Mountain Res. Dev.* **28**, 26 (2008).
3. C. J. Trac, A. H. Schmidt, S. Harrell, T. M. Hinckley, *Environ. Practice* **15**, 350 (2013).
4. W. F. Laurance *et al.*, *Science* **319**, 1184 (2008).

10.1126/science.aah4960

TECHNICAL COMMENT ABSTRACTS

Comment on "Mutation rate and genotype variation of Ebola virus from Mali case sequences"

Andrew Rambaut, Gytis Dudas, Luiz Max de Carvalho, Daniel J. Park, Nathan L. Yozwiak, Edward C. Holmes, Kristian G. Andersen

Hoenen *et al.* (Reports, 3 April 2015, p. 117; published online 26 March) suggested that

the Ebola virus Makona responsible for the West African epidemic evolved more slowly than previously reported. We show that this was based on corrupted data. An erratum provided a rate compatible with the initial and later, more precise, estimates but did not correctly state the nature of the error.

Full text at <http://dx.doi.org/10.1126/science.aaf3823>

Response to Comment on "Mutation rate and genotype variation of Ebola virus from Mali case sequences"

Thomas Hoenen, Allison Groseth, David Safronetz, Kurt Wollenberg, Heinz Feldmann

Rambaut *et al.* show that the erratum to our report on Ebola virus Makona evolution not only corrected sample dates modified by others in GenBank but also corrected an additional transcriptional error in our original analysis. We agree with their observation that both factors contributed to our revised evolutionary rate estimate but continue to stand by our revised estimate and conclusions.

Full text at <http://dx.doi.org/10.1126/science.aaf4561>

TECHNICAL COMMENT

VIOLOGY

Comment on “Mutation rate and genotype variation of Ebola virus from Mali case sequences”

Andrew Rambaut,^{1,2*} Gytis Dudas,¹ Luiz Max de Carvalho,¹ Daniel J. Park,³ Nathan L. Yozwiak,^{3,4} Edward C. Holmes,⁵ Kristian G. Andersen^{6,7,3*}

Hoenen *et al.* (Reports, 3 April 2015, p. 117; published online 26 March) suggested that the Ebola virus Makona responsible for the West African epidemic evolved more slowly than previously reported. We show that this was based on corrupted data. An erratum provided a rate compatible with the initial and later, more precise, estimates but did not correctly state the nature of the error.

In their Report, Hoenen and colleagues (1) presented an analysis of four Ebola virus (EBOV) genome sequences from Mali in the context of 102 previously published genomes from Guinea and Sierra Leone. Their key assertion was that the evolutionary rate of EBOV during the 2013 to 2016 West African Ebola virus disease (EVD) epidemic was lower than initially reported by Gire *et al.* (2) and similar to the long-term rate of evolution estimated over 35 years in a nonhuman reservoir, presumed to be a bat species. Hoenen *et al.* went on to state that this reported difference in evolutionary rate had important implications for the evolution of transmissibility and virulence of EBOV and for our attempts to control the EVD epidemic. We show that these conclusions are erroneous, based on the corruption of the data by the authors, and paint a false picture of EBOV evolution.

In an erratum to their paper, Hoenen and colleagues (3) substantially revised their evolutionary rate estimates and claimed that this was the consequence of changes to the sample collection dates of the Gire *et al.* (2) data set recorded in GenBank. However, these changes [which only affected 16 of the samples with no change in collection date of more than 6 days (table S1)] do not explain the low rate estimate provided in their original paper (Fig. 1). Hoenen *et al.* kindly

provided their original data file, and we confirm the apparently low evolutionary rate from these data (Fig. 1). However, comparing these sequences with those available on GenBank shows that an extensive shuffling of taxon labels (including date information) among sequences had occurred in their original publication.

This resulted in dates being assigned apparently randomly to sequences, with a subsequent loss of molecular clock signal. This is clearly evident in maximum likelihood phylogenetic trees estimated for both data sets; these possess identical topologies, but there is a clear mixing of label assignments (Fig. 2). As a case in point,

some sequences from the Gire *et al.* (2) data set were from sequential samples from the same patient taken a few days apart. Although these replicate sequences were mostly identical, they often occupy different positions in the Hoenen *et al.* phylogeny [figure 1 in (1)], confirming that a shuffling of sequences must have occurred.

The revised Report states, “By including the newly determined sequences from Mali, we obtained a mean substitution rate of 1.3×10^{-3} substitutions per site per year... This approaches previously reported nucleotide substitution rates of 0.6×10^{-3} to 1.0×10^{-3} for other EBOV sample sets... but is lower than the substitution rate of $\sim 1.9 \times 10^{-3}$ that had been reported for this outbreak.” Although we agree that a rate of 1.3×10^{-3} substitutions per site per year better reflects the evolutionary dynamics of EBOV during this outbreak, this revised estimate is between 1.4 and 2.1 times as high as the between-outbreak estimates cited by Hoenen *et al.* (1). However, the statistical credible interval for the revised Hoenen *et al.* (3) rate estimates is broad (Fig. 1), necessarily reflecting the limited time span and data at that point. The rate reported by Gire *et al.* (1.9×10^{-3}) was affected by the use of the three original Guinea sequences (4), which contained a number of imputation errors, later corrected (figure S2). These erroneous nucleotide sites were corrected for most of the analyses (as described in Gire *et al.*), but not for the rate value reported, and using the revised Guinea sequences brings the rate estimate for the Gire *et al.* data to 1.5×10^{-3} (figure S3). Importantly, both the revised Hoenen *et al.* (3) and Gire *et al.* (2) rate credible intervals are consistent with estimates reported by studies undertaken later in the epidemic (5–7) (Fig. 1).

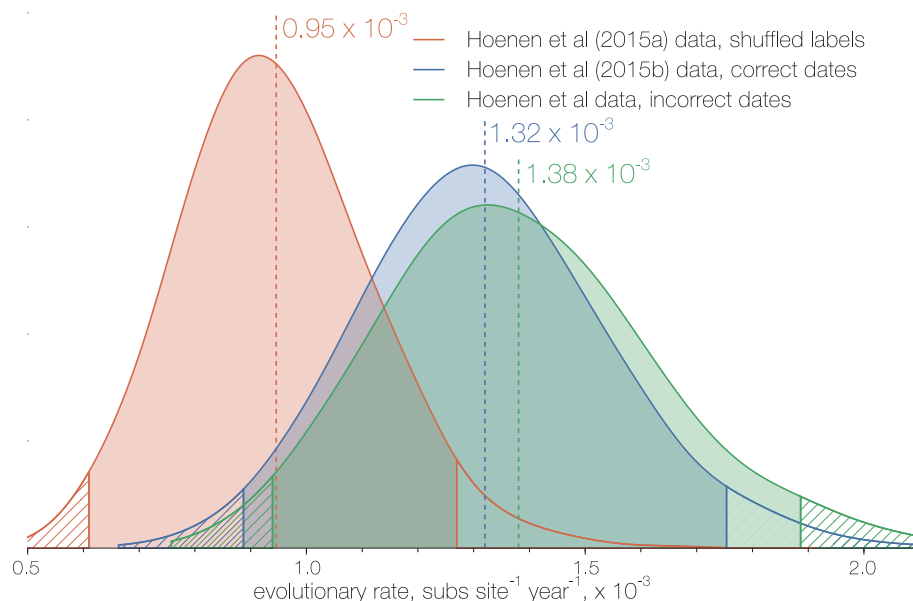


Fig. 1. Posterior probability densities of the mean evolutionary rate estimated for the sequence data provided by Hoenen *et al.*, including both sample shuffling and date errors, as originally published (red), sample date errors only (green), and as fully corrected in the erratum (blue). The vertical lines mark the 95% highest posterior density intervals.

¹Institute of Evolutionary Biology, University of Edinburgh, EH9 3FL, Edinburgh, UK. ²Centre for Immunology, Infection, and Evolution at the University of Edinburgh, EH9 3FL, Edinburgh, UK. ³Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. ⁴Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. ⁵Marie Bashir Institute of Infectious Diseases and Biosecurity, Charles Perkins Centre, Sydney Medical School, and School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, 2006, Australia. ⁶The Scripps Research Institute, Department of Immunology and Microbial Science, La Jolla, CA 92037, USA. ⁷Scripps Translational Science Institute, La Jolla, CA 92037, USA. *Corresponding author. Email: a.rambaut@ed.ac.uk (A.R.); kga1978@gmail.com (K.G.A.)

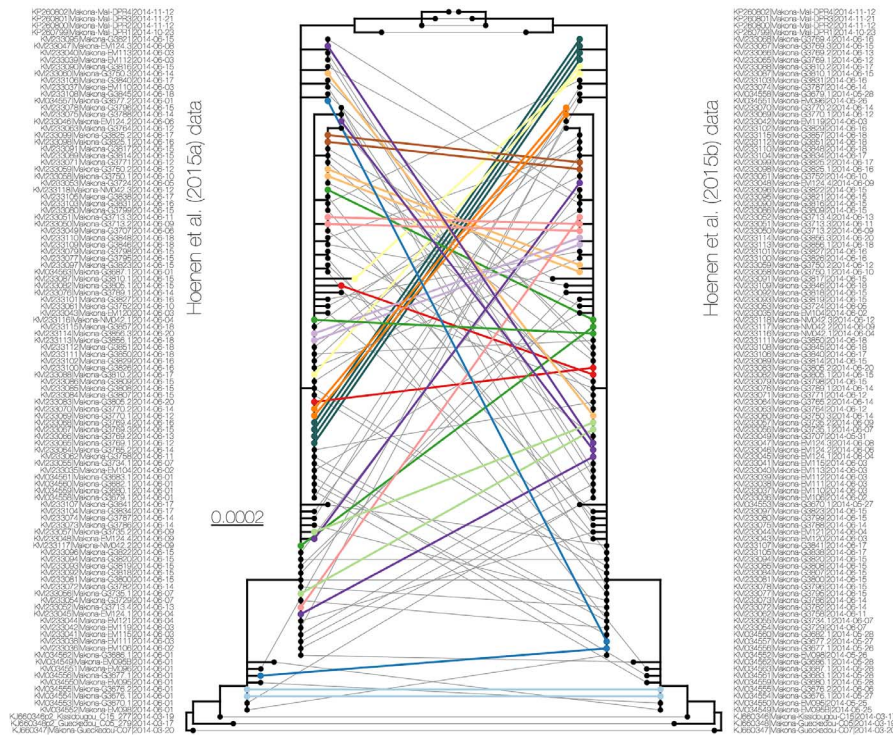


Fig. 2. Maximum likelihood tree of the 106 sequences analyzed by Hoenen *et al.* (left side) initially (1) with lines linking to the correctly labeled sequences after the erratum (3) (right side). Lines of the same color represent multiple samples taken from the same patient, which in most cases have identical sequences. These correctly group together on the right but do not in many cases on the left.

Based on their observed evolutionary rate estimate, Hoenen *et al.* (1) conclude that it is unlikely that the types of genetic changes observed thus far

would impair diagnostic measures or affect the efficacy of vaccines or potential virus-specific treatments. It is overly simplistic to conclude that the

differences in evolutionary rate estimates reported for EBOV will greatly alter its potential to change its virulence and/or transmissibility or our attempts to control it. For example, human dengue virus has been able to generate considerable antigenic diversity (8) that hinders successful vaccination, despite experiencing evolutionary rates usually lower than EBOV Makona (9). In addition, these reported differences in rate estimates are minor compared to the range of evolutionary rates seen in RNA viruses that span approximately 10^{-5} to 10^{-2} substitutions per site per year (9, 10). As underlying mutation rates in EBOV are still likely to be on the order of one mutation per genome replication, genotype variation is undoubtedly generated at a rate sufficient to enable rapid phenotypic evolution should a suitable selection pressure arise (11).

REFERENCES

1. T. Hoenen *et al.*, *Science* **348**, 117–119 (2015).
2. S. K. Gire *et al.*, *Science* **345**, 1369–1372 (2014).
3. T. Hoenen *et al.*, *Science* **348**, aac5674 (2015).
4. S. Baize *et al.*, *N. Engl. J. Med.* **371**, 1418–1425 (2014).
5. Y.-G. Tong *et al.*, *Nature* **524**, 93–96 (2015).
6. M. W. Carroll *et al.*, *Nature* **524**, 97–101 (2015).
7. D. J. Park *et al.*, *Cell* **161**, 1516–1526 (2015).
8. L. C. Katzelnick *et al.*, *Science* **349**, 1338–1343 (2015).
9. R. Sanjuán, *PLOS Pathog.* **8**, e1002685 (2012).
10. S. Duffy, L. A. Shackleton, E. C. Holmes, *Nat. Rev. Genet.* **9**, 267–276 (2008).
11. J. R. Kugelman *et al.*, *Cell Rep.* **12**, 2111–2120 (2015).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/aaf3823/suppl/DC1
Figs. S1 to S3

Table S1

References

5 February 2016; accepted 21 June 2016
10.1126/science.aaf3823

TECHNICAL RESPONSE

Virology

Response to Comment on “Mutation rate and genotype variation of Ebola virus from Mali case sequences”

Thomas Hoenen,¹ Allison Groseth,¹ David Safronetz,¹
Kurt Wollenberg,² Heinz Feldmann^{1*}

Rambaut *et al.* show that the erratum to our report on Ebola virus Makona evolution not only corrected sample dates modified by others in GenBank but also corrected an additional transcriptional error in our original analysis. We agree with their observation that both factors contributed to our revised evolutionary rate estimate but continue to stand by our revised estimate and conclusions.

In March of 2015, we reported several new sequences from the recent West African Ebola virus (EBOV) Makona epidemic, along with a phylogenetic analysis of these and all other sequences available in GenBank at that time (1). Upon request, we immediately shared our data (including detailed analysis files) with Rambaut and colleagues, who had coauthored an earlier study, which had reported that EBOV in the West African epidemic was mutating roughly twice as fast as historically observed (2). However, at the same time, we also noticed that the sampling dates for 15 samples in GenBank had been changed. Given that this represented 14% of the data set, and because almost all of the samples in this data set were collected during a very short time frame of 26 days, we were concerned that these changes of up to 6 days might alter the outcome of our analysis. This prompted us to redo our whole analysis. We published the results of this second analysis in an erratum in May 2015 (3).

The rigorous analysis presented here by Rambaut *et al.* (4) using the data set we shared with them, and which was used to generate our original evolutionary rate estimate, has now, as is clearly shown in their Technical Comment, uncovered a further error in this original data set. Unfortunately, and unbeknownst to us, when transcribing the sampling dates of the genome sequences for phylogenetic analysis, a mistake occurred—specifically, one block of sampling dates was inadvertently shifted relative to the sample names. As a consequence of repeating our analysis using a data set completely reassembled from scratch for the erratum, this additional transcription error was

also corrected at the same time. The analysis by Rambaut *et al.*, although clearly supporting our revised evolutionary rate estimate (3), now shows that the correction of this transcription error during our reanalysis of the data was primarily responsible for the difference between our original and revised estimates.

Rambaut *et al.* are indeed correct in their analyses: A mistake was made when transcribing some of the original data for the analysis software, which of course we deeply regret. However, both this mistake and the revisions to the primary data in GenBank have already been corrected in our erratum last year. Importantly, the revised evolutionary rate estimate derived from that analysis, and reported in our erratum, remains correct and has now been confirmed by several other more recent studies that benefited from more diversified data sets based on longer-term sampling efforts (5–7). Further, in their Comment, Rambaut *et al.* now also point out that the value of 1.9×10^{-3} that they originally reported was based on erroneous data and that a revised analysis by them using corrected data gives a rate estimate coinciding with our revised rate estimate, as well as more recent estimates from other groups, so that there is now a general consensus in the field as to this rate.

Regarding the main conclusions of our study, we also believe that these remain correct. We are not aware of any epidemiological evidence of changes in transmission (something that remains true even considering recently demonstrated rare late sexual transmissions from survivors). Overall, case fatality rates in this epidemic appear to have been somewhat lower than during previous outbreaks, even though confounding factors (particularly the extent of diagnostics performed) complicate such a comparison. This also seems supported by initial infection studies in macaques, which did not result in increased virulence of the West African EBOV Makona strain (8), as well as by the first experiments in humanized mice, where infection with a wild-type EBOV

Makona isolate resulted in a longer mean-time-to-death than infection with the prototype EBOV Mayinga strain from 1976 (9). Similarly, a vaccine that was developed against the EBOV Kikwit strain from 1995 remains protective against the West African EBOV Makona strain from 20 years later in the macaque model (10) and has shown promising results in a human ring vaccination trial (11). Also, a first experimental study addressing sequence changes occurring early in the epidemic did not reveal evidence for altered virus biology between different EBOV Makona variants or between EBOV Makona and EBOV Mayinga (12), and comparative studies of virus entry also revealed no differences between EBOV Makona and EBOV Mayinga (13). Lastly, so far as we are aware, laboratory diagnostics have not been affected by sequence changes over the period of the epidemic, which is of course one of the biggest concerns in outbreak response, and systematic in vitro evaluation of available diagnostic reverse transcription polymerase chain reaction assays showed that they perform virtually identically for the EBOV Makona and the EBOV Kikwit strains, which were isolated 20 years apart (14). Overall, we believe that our study has directly helped to balance some misconceptions in the media and public about this emerging EBOV virus strain and thus supported proper public health intervention.

In summary, Rambaut and colleagues have indeed identified a technical error in our analysis, which had already been corrected in our erratum last year, and we certainly appreciate their efforts to further correct the scientific record by identifying this error. However, we continue to stand by our revised evolutionary rate estimate, and believe that the major conclusions drawn from it have also been well supported by the experiences made during this epidemic.

REFERENCES AND NOTES

1. T. Hoenen *et al.*, *Science* **348**, 117–119 (2015).
2. S. K. Gire *et al.*, *Science* **345**, 1369–1372 (2014).
3. T. Hoenen *et al.*, *Science* **348**, aac5674 (2015).
4. A. Rambaut *et al.*, *Science* **353**, 658 (2016).
5. M. W. Carroll *et al.*, *Nature* **524**, 97–101 (2015).
6. D. J. Park *et al.*, *Cell* **161**, 1516–1526 (2015).
7. Y. G. Tong *et al.*, *Nature* **524**, 93–96 (2015).
8. A. Marzi *et al.*, *Emerg. Infect. Dis.* **21**, 1777–1783 (2015).
9. B. H. Bird *et al.*, *J. Infect. Dis.* **213**, 703–711 (2016).
10. A. Marzi *et al.*, *Science* **349**, 739–742 (2015).
11. A. M. Henao-Restrepo *et al.*, *Lancet* **386**, 857–866 (2015).
12. E. C. Dunham *et al.*, *Nat. Commun.* **6**, 8000 (2015).
13. H. Hofmann-Winkler, K. Gnirß, F. Wrensch, S. Pöhlmann, *J. Infect. Dis.* **212** (suppl. 2), S172–S180 (2015).
14. S. Sozhamannan *et al.*, *Viruses* **7**, 3130–3154 (2015).

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of NIH, National Institute of Allergy and Infectious Diseases.

25 February 2016; accepted 21 June 2016
10.1126/science.aaf4561

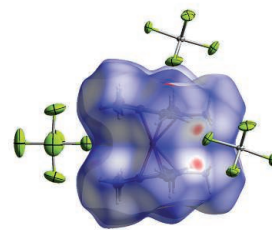
¹Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, MT, USA. ²Bioinformatics and Computational Biosciences Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA.

*Corresponding author. Email: feldmannh@niaid.nih.gov

RESEARCH

Molecular structure of a metallocene

Malischewski et al., p. 678



IN SCIENCE JOURNALS

Edited by **Caroline Ash**



HYDROLOGY

Water dissolving and water removing

Not all groundwater ends up flowing into rivers. Some is discharged directly into the ocean along the coast. Although much lower in volume than water transported by rivers, such submarine groundwater discharge can be a hidden source of dissolved ions, nutrients, or contaminants from human activities. Sawyer *et al.* performed a high-resolution continental-scale analysis of fresh groundwater discharge along the coastline of the United States. In total, more than one-fifth of coastal waters are vulnerable to groundwater-borne contamination. —NW

Science, this issue p. 705

Seepage of coastal groundwater rich in iron and other dissolved constituents

PHYSICS

The deuteron is too small, too

The radius of the proton has remained a point of debate ever since the spectroscopy of muonic hydrogen indicated a large discrepancy from the previously accepted value. Pohl *et al.* add an important clue for solving this so-called proton radius puzzle. They determined the charge radius of the deuteron, a nucleus consisting of a proton and a neutron, from the transition frequencies in muonic deuterium. Mirroring the proton radius puzzle, the radius of the deuteron was several standard deviations smaller than the value inferred from previous spectroscopic measurements of electronic deuterium. This

independent discrepancy points to experimental or theoretical error or even to physics beyond the standard model. —JS

Science, this issue p. 669

LIFE HISTORY

Deep living for centuries

We tend to think of vertebrates as living about as long as we do, give or take 50 to 100 years. Marine species are likely to be very long-lived, but determining their age is particularly difficult. Nielsen *et al.* used the pulse of carbon-14 produced by nuclear tests in the 1950s—specifically, its incorporation into the eye during development—to determine the age of Greenland sharks. This species is large yet slow-growing. The oldest of the animals that they sampled had

lived for nearly 400 years, and they conclude that the species reaches maturity at about 150 years of age. —SNV

Science, this issue p. 702

BRAIN MICROCIRCUITS

Building new networks in the brain

Donald Hebb's hypothesis that coactivation of neurons leads to the formation of ensembles of neurons has inspired neuroscientists for decades. The experimental creation of such ensembles has been technically challenging. Using two-photon optogenetic stimulation with single-cell resolution, Carrillo-Reid *et al.* discovered that recurrent activation of a group of neurons creates an ensemble

that is imprinted in the brain circuitry. Activation of a single neuron can lead to recall of the entire ensemble in a phenomenon called pattern completion. The artificial ensemble persists over days and can be reactivated at later time points without interfering with endogenous circuitry. —PRS

Science, this issue p. 691

ATMOSPHERIC CHEMISTRY

Active fatty acid layers

Saturated fatty acids are considered to be inert, but they can be surprisingly reactive when present as a coating at an air-water interface. Rossignol *et al.* show that nonanoic acid is photochemically active when it is present as a monolayer on a water surface (see the

Perspective by Vaida). Fatty acids are ubiquitous in the environment, and their photochemical processing could have a substantial impact on local ozone and particle formation. —HJS

Science, this issue p. 699;
see also p. 650

ECONOMIC POLICY

Programs that buffer a financial shock work

For people without a safety net of social and financial resources, a shock, such as medical expenses not covered by insurance, can be the first step in a downward spiral toward homelessness and morbidity. Evans *et al.* evaluate the effectiveness and cost of a program in Chicago that provides temporary financial assistance with the aim of enabling individuals to stay in their homes and out of homeless shelters. They find that one-time payments of up to \$1500 greatly reduce the likelihood of homelessness. The estimated economic benefits exceed the estimated costs, with immeasurable psychic and physical benefits. —GJC

Science, this issue p. 694

SOFT ELECTRONICS

Soft and still responsive

Transparent touch screens, from large-panel interactive information maps to advanced cell phones, have become a part of daily life. However, such devices all use hard materials. Kim *et al.* have developed a soft touch panel based on



Soft touch-conductive hydrogel

polyacrylamide hydrogels (cross-linked polymers swollen with water) that are highly transparent and contain trapped LiCl to enhance conductivity. The hydrogels are soft and can be stretched extensively while still maintaining touch sensitivity. —MSL

Science, this issue p. 682

BONE DEVELOPMENT

Turning chondrocytes into bone killers

The skeletal defects caused by inhibitors of histone deacetylase (HDAC) enzymes limit the clinical value of these drugs. Carpio *et al.* found that a specific isoform, HDAC3, promotes bone growth by restricting the secretion of inflammatory factors from cartilage cells called chondrocytes. Mice that lacked *Hdac3* in chondrocytes after birth had impaired long bone development. Chondrocytes from these mice had increased activation of a proinflammatory transcription factor. The findings help to explain why HDAC inhibitors are not a good option for children and pregnant women and for patients with bone fractures. —LKF

Sci. Signal. **9**, ra79 (2016).

TRANSPLANTATION

Make way for stem cells

Current chemotherapy or radiation regimens to prepare the host bone marrow for transplantation of donor hematopoietic stem cells (HSCs) can be harmful. Chhabra *et al.* tested an alternative strategy in which the surface antigen CD47 is blocked, which allows phagocytic myeloid cells to engulf host HSCs that are displaced by antibody targeting, effectively depleting HSCs from the bone marrow of immunocompetent mice and enabling engraftment of donor cells. —LP

Sci. Transl. Med. **8**, 351ra105 (2016).

IN OTHER JOURNALS

Edited by **Kristen Mueller**
and **Jesse Smith**



Symbiotic root fungi helps trees access nutrient 'patches' in soil.

SYMBIOSIS

Fungi help trees hunt for food

Trees face a difficult paradox: how to access nutrients that are not uniformly spread throughout the soil while remaining stationary. Nearly all plant roots associate with symbiotic soil-dwelling fungi (either intracellular arbuscular mycorrhizal or extracellular ectomycorrhizal fungi), which aid in nutrient uptake. Chen *et al.* now report that mycorrhizae help trees forage. Tree species with finer roots that associate with arbuscular mycorrhizal fungi produce more roots when they encounter a nutrient patch, and those that associate with ectomycorrhizal fungi produce more fungal hyphae. Moreover, trees in mixed woodlands probably have complementary foraging strategies by virtue of their differing symbionts, likely contributing to tree diversity in temperate forests. —CA

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1601006113 (2016).

PSYCHOLOGY

The persuasiveness of reductionism

A decade ago, it seemed as though every other neuroscience paper in high-profile journals featured multiple multicolored images of brain scans.

In some cases, readers—many of whom were psychologists who had written papers on the same topic—pointed out that the pictographic scans added little explanatory power. Hopkins *et al.* have extended an earlier study of the relative impact of psychology and

ALSO IN SCIENCE JOURNALS

Edited by Caroline Ash

ZIKA

Global spread of Zika virus

Zika virus was identified in Uganda in 1947; since then, it has enveloped the tropics, causing disease of varying severity. Lessler *et al.* review the historical literature to remind us that Zika's neurotropism was observed in mice even before clinical case reports in Nigeria in 1953. What determines the clinical manifestations; how local conditions, vectors, genetics, and wild hosts affect transmission and geographical spread; what the best control strategy is; and how to develop effective drugs, vaccines, and diagnostics are all critical questions that are begging for data. —CA

Science, this issue p. 663

STRUCTURAL BIOLOGY

A different gate design

The voltage-gated potassium channel Eag1 is overexpressed in tumor cells from a range of cancers, and inhibiting Eag1 reduces tumor growth. Whicher and Mackinnon determined the structure of a mammalian Eag1 bound to the inhibitor calmodulin at 3.78 Å resolution (see the Perspective by Toombes and Swartz). The organization of the voltage-sensing and pore domains differs from that of other potassium channels, indicating that the gating mechanism is distinct. The structure also shows how the channel can be closed by a ligand, independently of the position of the voltage sensor. —VV

Science, this issue p. 664;
see also p. 646

EXTRASOLAR PLANETS

Spying a planet in a triple-star system

Thousands of extrasolar planets are now known, but only a handful have been detected in direct images. Wagner *et al.* used

sophisticated adaptive optics to discover a planet in images of the triple-star system HD 131399 and to take a spectrum of its atmosphere (see the Perspective by Oppenheimer). The planet, about four times the mass of Jupiter, orbits around one star in the system while the other two stars move farther out. This unusual arrangement is puzzling: The planet's orbit may be stable, but it is unclear how it could have formed or migrated there. The results will be used to refine theories of planet formation. —KTS

Science, this issue p. 673;
see also p. 644

SLEEP RESEARCH

Circadian rhythms and sleep deprivation

Sleep deprivation, such as that experienced because of shift work, jet lag, sleep disorders, and aging, leads to deterioration of many aspects of health. Cognition deteriorates rapidly and substantially when we stay awake through the night. To investigate the time course of brain responses during sleep loss, Muto *et al.* scanned volunteers repeatedly during an extended period of wakefulness (see the Perspective by Czeisler) in which circadian and homeostatic drives differentially affected local brain regions. —PRS

Science, this issue p. 687;
see also p. 648

NEURODEGENERATION

Targeting three defects with one strategy

The neurodegenerative diseases amyotrophic lateral sclerosis and frontotemporal dementia are most commonly caused by a mutation in the *C9orf72* gene. The mutation is an expanded hexanucleotide repeat in a noncoding region. The expanded repeat produces sense and

antisense RNA transcripts, which accumulate in patient cells and appear to sequester RNA-binding proteins. The sense and antisense transcripts are also translated into dipeptide repeat proteins, which are aggregation-prone and accumulate in the brain and spinal cord. Last, loss of function from reduced expression of *C9orf72* in neurons and glia could contribute to the disease. Kramer *et al.* targeted both sense and antisense repeats by blocking a single gene called *SPT4*, which mitigated degeneration in human cells by reducing all three types of pathologies. —SMH

Science, this issue p. 708

CONSERVATION

For richer, for poorer

Human activities are increasingly causing ecosystem degradation. In a Perspective, Frank and Schlenker discuss some of the trade-offs between economic development and conservation. They argue that although it can be difficult to assign monetary value to ecosystems, not doing so risks giving them zero value. For effective conservation, governments must either send the right price signals or place ecosystems directly under protection. Between 1972 and 2012, richer countries have seen the largest increase in protected areas, showing that economic growth can be compatible with conservation. —JFU

Science, this issue p. 651

ALZHEIMER'S DISEASE

Mapping Alzheimer's disease

Alzheimer's disease (AD) is a chronic neurodegenerative disorder that progresses slowly into dementia and short-term memory loss. AD is thought to be caused by aggregation of Ab and tau proteins in plaques and tangles. Vendruscolo *et al.* investigated why aberrant

protein aggregations are found in some tissues but not in others. It turns out that failure of protein homeostasis elevates expression of a specific subset of proteins that can be aggregated in AD. Thus, a tissue vulnerability map based on Ab and tau homeostasis in healthy tissue might be predictive for AD. —ASH

Sci. Adv. 10.1126/sciadv.1600947
(2016).

ORGANOMETALLICS

Charging up the iron in ferrocene salts

Ferrocene is the archetype of the sandwich compounds, so called because a metal atom is inserted between two carbon rings. The elucidation of ferrocene's structure was pivotal to the development of organometallic chemistry during the mid-20th century. The ease with which the iron in the center of the molecule can toggle between the +2 and +3 oxidation states has made the compound a common electrochemical standard. Malischewski *et al.* report the synthesis and isolation of ferrocene salts with iron in the +4 state, which they characterize crystallographically and spectroscopically. —JSY

Science, this issue p. 678



NEAR-FIELD COSMOLOGY

Is the Local Group useful for cosmology?

Our Milky Way Galaxy is the second largest member of the Local Group, a gravitationally bound assemblage of several dozen galaxies. Researchers in near-field cosmology use detailed studies of the Local Group to understand how galaxies form and evolve over cosmic time but have never been sure how well the Local Group represents the wider universe. Boylan-Kolchin *et al.* used cosmological simulations to demonstrate that the Local Group occupies a sufficiently large volume at high redshift to make it a representative sample, at least for the dwarf galaxies. They show that Local Group studies can complement deep observations of high-redshift galaxies, which are sensitive to different mass ranges. —KTS

Mon. Not. R. Astron. Soc. **462**, L51 (2016).

A nighttime view from Earth of our Milky Way Galaxy

neuroscience to encompass both more reductive disciplines, such as physics, chemistry, and biology, and less reductive disciplines, such as social science. They find that study subjects judge scientific explanations to be of higher quality when they contain information from the neighboring more reductive field, even when that information is irrelevant. —GJC

Cognition **155**, 67 (2016).

MAGNETISM

Measuring the elusive interaction

When two spins interact, the part of the interaction that changes sign when the spins are exchanged is called the Dzyaloshinsky-Moriya (DM) interaction. The DM interaction favors nonparallel spins and can play an important role in the formation of skyrmions (miniature whirlpools of spins) and in the magnetoelectric effect in multiferroics. However, measuring both the size and the direction of this interaction is tricky. Laplane *et al.* conceived a general method that they demonstrated on a pair of Nd³⁺ ions embedded as dopants in a YVO₄ crystal. Working at low temperatures where the Nd³⁺

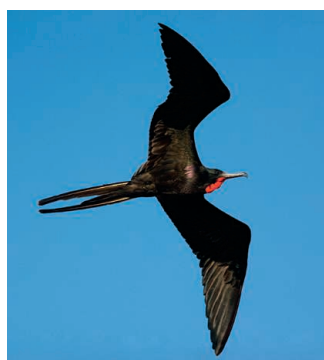
ions behaved as effective $\frac{1}{2}$ spins, they extracted the details of the interaction by using the electron spin resonance technique combined with optical detection. —JS

Phys. Rev. Lett. **117**, 037203 (2016).

ECOPHYSIOLOGY

Staying away for the long haul

Frigatebirds can fly without stopping for months. Such long flights, however, present considerable challenges to some physiological processes, notably sleep. Scientists think that birds undergoing long flights use hemispheric sleep, where half of the brain sleeps at a time. To find



Frigatebirds can fly long distances with little need for sleep.

out whether this is indeed the case, Rattenborg *et al.* placed mobile electroencephalogram recorders on flying frigatebirds and found that although they do use hemispheric sleep, especially when riding updrafts, they actually sleep remarkably little during their long flights. Though the birds may be able to catch up on their sleep when on land, it seems that they can mostly avoid the sleep deprivation effects that plague most vertebrates, an ability that is probably shaped by strong selection for wakefulness during flight. —SNV

Nat. Comm. **10**, 1038/ncomms12468 (2016).

CLIMATE WARMING

Warming our world

How much will human emissions of carbon dioxide cause global temperatures to rise? The magnitude of that warming depends a great deal on the response of clouds: If clouds reflect more sunlight back into space, warming will be reduced, but if they reflect less, warming will be greater. Brient and Schneider use satellite data to show that low clouds over tropical oceans will reflect less shortwave radiation as surface waters warm, supporting estimates of climate

sensitivity to atmospheric carbon dioxide concentrations on the higher end of the existing range, and that this behavior will make it unlikely that global temperature rise can be capped at less than the common target of 2.0°C. —HJS

J. Clim. **10**, 1175/

JCLI-D-15-0897.1 (2016).

RNA STABILITY

Codon optimality at genome transition

Nucleotide triplets, or codons, designate specific amino acids for protein synthesis. However, that is not their only job. In yeast and bacteria, codons contribute to RNA stability, with “optimal” codons stabilizing RNAs and “suboptimal” codons destabilizing RNAs. This is possible because multiple codons can encode the same amino acid. Bazzini *et al.* now demonstrate that codon usage in zebrafish, frogs, mice, and flies can affect transcript degradation and polyadenylation at the critical stage in development when transcription switches from the maternal to the zygotic genome. Furthermore, enriching genes with nonoptimal codons can reduce translation efficiency. —BAP

EMBO J. **10**, 15252/

embj.201694699 (2016).

PHOTOS (FROM TOP): BABAK TA'FRESHI/NATIONAL GEOGRAPHIC CREATIVE; BOB GIBBONS/ALAMY STOCK PHOTO

REVIEW SUMMARY

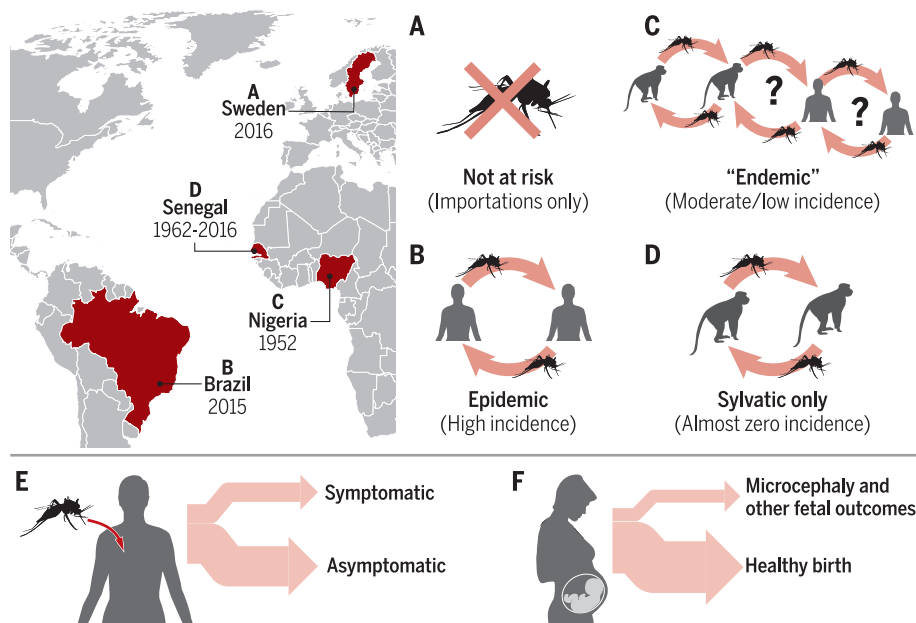
ZIKA

Assessing the global threat from Zika virus

Justin Lessler,^{*†} Lelia H. Chaisson,[†] Lauren M. Kucirka, Qifang Bi, Kyra Grantz, Henrik Salje, Andrea C. Carcelen, Cassandra T. Ott, Jeanne S. Sheffield, Neil M. Ferguson, Derek A. T. Cummings, C. Jessica E. Metcalf, Isabel Rodriguez-Barraquer

BACKGROUND: First discovered in 1947, Zika virus (ZIKV) received little attention until a surge in microcephaly cases was reported after a 2015 outbreak in Brazil. The size of the outbreak and the severity of associated birth defects prompted the World Health Organization (WHO) to declare a Public Health Emergency of International Concern on 1 February 2016. In response, there has been an explosion in research and planning as the global health community has turned its attention to understanding and controlling ZIKV. Still, much of the information needed to evaluate the global health threat from

ZIKV is lacking. The global threat posed by any emerging pathogen depends on its epidemiology, its clinical features, and our ability to implement effective control measures. Whether introductions of ZIKV result in epidemics depends on local ecology, population immunity, regional demographics, and, to no small degree, random chance. The same factors determine whether the virus will establish itself as an endemic disease. The burden of ZIKV spread on human health is mediated by its natural history and pathogenesis, particularly during pregnancy, and our ability to control the virus's spread. In this Review, we examine the



The effect of ZIKV is a function of the local transmission regime and viral pathogenesis.

(A) Many countries cannot maintain ongoing vector-mediated ZIKV transmission and are only at risk from importation by travelers and limited onward transmission (e.g., through sex). (B) If conditions are appropriate, importations can lead to postinvasion epidemics with high incidence across age ranges, after which the virus may go locally extinct or remain endemic. (C) There is evidence of ongoing ZIKV incidence in humans over years (e.g., a 1952 serosurvey in Nigeria), but it is unknown whether this is the result of ongoing circulation in humans or frequent spillover infections from a sylvatic cycle. (D) In other areas, ZIKV appears to have been maintained in animals with few human infections. (E) The majority of infections are asymptomatic, and severe outcomes, such as Guillain-Barré syndrome, are rare. (F) However, there is considerable risk of microcephaly and other fetal sequelae when infection occurs during pregnancy.

empirical evidence for a global threat from ZIKV through the lens of these processes, examining historic and current evidence, as well as parallel processes in closely related viruses.

ADVANCES: Because ZIKV was not recognized as an important disease in humans until recently, it was little studied before the recent crisis. Nevertheless, the limited data from the decades following its discovery provide important clues into ZIKV's epidemiology and suggest that some populations were at risk for the virus for years in the mid-20th century, although this risk may predominantly have been the result of spillover infections from a sylvatic reservoir.

ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aaf8160>

Recent outbreaks on Yap Island (2007) and in French Polynesia (2014) provide the only previous observations of large epidemics and are the basis for the little that we do know about ZIKV's acute symptoms (e.g., rash, fever, conjunctivitis, and arthralgia), the risk of birth defects, such as microcephaly (estimated to be 1 per 100 in French Polynesia), and the incidence of severe neurological outcomes (e.g., Guillain-Barré is estimated to occur in approximately 2 out of every 10,000 cases). The observation of an association between ZIKV and a surge in microcephaly cases in Brazil and the subsequent declaration of a Public Health Emergency of International Concern by the WHO have rapidly accelerated research into the virus. Small, but very important, studies have begun to identify the substantial risk the virus can pose throughout a pregnancy, and careful surveillance has established that ZIKV can be transmitted sexually. Numerous modeling studies have helped to estimate the potential range of ZIKV and measured its reproductive number R_0 (estimates range from 1.4 to 6.6), a key measure of transmissibility in a number of settings. Still, it remains unclear whether the recent epidemic in the Americas is the result of fundamental changes in the virus or merely a chance event.

OUTLOOK: ZIKV research is progressing rapidly, and over the coming months and years our understanding of the virus will undoubtedly deepen considerably. Key questions about the virus's range, its ability to persist, and its clinical severity will be answered as the current epidemic in the Americas runs its course. Moving forward, it is important that information on ZIKV be placed within the context of its effect on human health and that we remain cognizant of the structure of postinvasion epidemic dynamics as we respond to this emerging threat. ■

The list of author affiliations is available in the full article online.

^{*}Corresponding author. Email: justin@jhu.edu

[†]These authors contributed equally to this work.

Cite this article as J. Lessler et al., *Science* 353, aaf8160 (2016). DOI: 10.1126/science.aaf8160

REVIEW

ZIKA

Assessing the global threat from Zika virus

Justin Lessler,^{1,*} Lelia H. Chaisson,^{1,†} Lauren M. Kucirka,^{1,2} Qifang Bi,¹ Kyra Grantz,³ Henrik Salje,^{1,4} Andrea C. Carcelen,⁵ Cassandra T. Ott,¹ Jeanne S. Sheffield,⁶ Neil M. Ferguson,⁷ Derek A. T. Cummings,³ C. Jessica E. Metcalf,^{8,9} Isabel Rodriguez-Barraquer¹

First discovered in 1947, Zika virus (ZIKV) infection remained a little-known tropical disease until 2015, when its apparent association with a considerable increase in the incidence of microcephaly in Brazil raised alarms worldwide. There is limited information on the key factors that determine the extent of the global threat from ZIKV infection and resulting complications. Here, we review what is known about the epidemiology, natural history, and public health effects of ZIKV infection, the empirical basis for this knowledge, and the critical knowledge gaps that need to be filled.

Originally discovered in 1947, Zika virus (ZIKV) received little attention until a surge in microcephaly cases was reported after a 2015 outbreak in Brazil (1, 2). Prompted by the size of the outbreak and the severity of associated birth defects, the World Health Organization (WHO) declared ZIKV to be a Public Health Emergency of International Concern on 1 February 2016 (3). In response, there has been an explosion in research and planning as the global health community has turned its attention to understanding and controlling ZIKV. Still, much of the information needed to evaluate the global health threat from ZIKV remains unknown.

The global threat posed by any emerging pathogen depends on its epidemiology, its clinical features, and our ability to implement effective control measures (Fig. 1). In an interconnected world, introductions of ZIKV to areas free of the virus may be inevitable. Whether these introductions result in only a few subsequent cases or a major epidemic depends on the local ecology, population immunity, demographics of the region, and random chance. The ability of the virus to transmit in any area can be characterized by its reproductive number R : the number of people we expect to become infected from each case in

that area (4). When R is greater than one, an epidemic can occur, and when it is less than one, onward transmission will be limited. When ZIKV successfully invades, the threat may be transient and the virus might become locally extinct, as appears to have been the case in Yap Island and French Polynesia (5, 6), or it may persist endemically, as seems to be the case in parts of Africa (7). There are two ways in which ZIKV can persist in a region: through ongoing transmission in animals (i.e., a sylvatic cycle) with occasional spillover into the human population, or through sustained transmission in humans (8, 9). Which-ever scenario emerges, the natural history and pathogenesis of ZIKV will determine its effect on human health, with infection in pregnant women being particularly important (10). Finally, the extent of the global threat from ZIKV is mediated by our ability to control the virus and treat those cases that do occur.

In this review, we examine the empirical evidence for a global threat from ZIKV through the lens of these processes. We review what is known about the natural history and pathogenesis of ZIKV in humans, outline what we know about the ability of ZIKV and similar viruses to invade and persist in diverse settings, and summarize the challenges we face in studying and controlling ZIKV. Finally, we examine what we know about why ZIKV has emerged as a public health threat in the Americas after being known for decades as a rare and mild tropical disease.

A brief history of ZIKV

ZIKV was discovered in the blood of a rhesus monkey in 1947 at the Yellow Fever Research Institute in Entebbe, Uganda (1), and was isolated from *Aedes africanus* mosquitoes the following year (1). Soon after, multiple serosurveys found evidence of antibodies to ZIKV in human populations throughout Africa (11–14), India (15), and Southeast Asia (16, 17) (Fig. 2). It was not initially clear that ZIKV caused clinical disease (13), al-

though early evidence suggested that it was neurotropic in mice (18). Human infection was first confirmed in 1953 in Nigeria (13), and ZIKV was definitively established as pathogenic in humans after later experimental (19) and natural (20) infections led to symptoms of fever and rash.

The globally distributed mosquito *A. aegypti* was identified as a likely vector for ZIKV transmission in the 1950s after successful transmission of the virus to a mosquito from an infected human volunteer (19). Later experiments confirmed *A. aegypti*'s ability to transmit ZIKV to mice (21), and ZIKV has since been isolated from several *Aedes* species (and, in a few cases, other genera) (22), including *A. albopictus* (23–26).

In the decades after its discovery, intermittent serosurveys continued to find evidence of ZIKV infection in humans in Africa (27–29), the Indian subcontinent (30), and Southeast Asia (16, 31, 32). Evidence for ZIKV's continued presence was further bolstered by limited viral isolations from mosquitoes (33–38), humans (7, 20, 29, 39, 40), and nonhuman primates (9). However, few clinical cases had been reported in humans before 2007 (20, 29, 31, 40), and ZIKV was considered to be of limited public health importance.

In 2007, the first known major outbreak of ZIKV occurred on Yap Island in the Federated States of Micronesia (6). Although several patients initially tested positive for dengue, the unusual clinical presentation prompted physicians to send serum to the Centers for Disease Control and Prevention (CDC) Arbovirus Diagnostic and Reference Laboratory, where it tested positive for ZIKV (6, 41). During the outbreak, ~73% of the island's residents were infected with ZIKV, and symptoms were generally mild and short-lived (6).

After the Yap Island outbreak, there were sporadic isolations of ZIKV in residents of and travelers to Southeast Asia (42–44), but no other major ZIKV outbreaks were observed until late 2013. From October 2013 to April 2014, French Polynesia experienced a large outbreak of ZIKV, estimated to have infected 66% of the general population (5, 45). A contemporaneous surge in the number of cases of Guillain-Barré syndrome raised concerns of an association with ZIKV (5, 45): A total of 42 cases of Guillain-Barré syndrome were reported from November 2013 to February 2014, compared with three cases in all of 2012. These are the first known instances of neurologic sequelae associated with ZIKV infection. Although not noted at the time, retrospective analyses suggest that there may also have been an increase in microcephaly cases (46). After the French Polynesia outbreak, ZIKV spread throughout the South Pacific, including outbreaks in New Caledonia, the Cook Islands, and Easter Island in 2014 (47).

The earliest confirmed cases of ZIKV infection in the Americas occurred in late 2014 in northeastern Brazil (48). Recent work suggests that the virus may also have been present simultaneously in Haiti (49). Over the following months, the virus spread rapidly throughout Brazil (50), followed by a substantial rise in cases of Guillain-Barré syndrome and microcephaly in affected regions (51), prompting

¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. ²Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ³Department of Biology, Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA. ⁴Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France. ⁵Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. ⁶Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁷Department of Medicine, School of Public Health, Imperial College London, London, UK. ⁸Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. ⁹Office of Population Research, Princeton University, Princeton, NJ, USA. *Corresponding author. Email: justin@jhu.edu †These authors contributed equally to this work.

the WHO to declare a Public Health Emergency of International Concern on 1 February 2016 (3). Phylogenetic evidence suggests that the strains that seeded this outbreak are descendants of those that caused outbreaks in the South Pacific, which in turn descended from the Asian lineage of the virus (52).

Since late 2014, ZIKV has spread widely throughout South and Central America and the Caribbean (2). As of June 2016, more than 35 countries throughout the Americas have reported locally circulating ZIKV (53). This includes a large outbreak in Colombia, with more than 65,000 reported cases, numerous reports of potentially associated neurological syndromes, and ZIKV-associated microcephaly cases (54–56). As of June 2016, the ZIKV situation continues to evolve, and the global threat ultimately posed by ZIKV remains uncertain.

The natural history and pathogenesis of ZIKV

Transmission and natural history of ZIKV

The primary source of ZIKV infection in humans is from bites of infected mosquitoes (57), although there have also been cases of sexual (58–60), perinatal (61), and suspected blood-transfusion transmission (62). Evidence from outbreaks in the South Pacific indicates that a minority of those infected with ZIKV develop clinical illness: During the Yap Island outbreak, 19% of people with serological evidence of recent infection [immunoglobulin M (IgM)-positive] reported ZIKV symptoms (6); in French Polynesia, 26% of ZIKV-positive blood donors who were asymptomatic at the time of donation later reported symptoms (63).

On average, those who do develop ZIKV symptoms will do so 6 days after infection (64), and 95% will do so within 11 days (Fig. 3). Virus has been isolated from blood (13), urine (65, 66), saliva (67), semen (68), amniotic fluid (69), and neurologic tissue (70). Virus can be isolated in blood for an average of 10 days after infection (99% will clear the virus by 24 days) (64), and case reports indicate that virus may remain in urine for 12 or more days after infection (65) and in semen for more than 60 days (59). Pregnancy may affect the length of viral shedding: In one case, a woman remained viremic for at least 10 weeks during pregnancy but cleared the virus within 11 days of termination (71). Antibodies to ZIKV become detectable, on average, 9 days after infection (64). Although the duration of immunity against ZIKV remains unknown, evidence from other flaviviruses suggests that it should be life-long (72). Mosquitoes become infectious about 10 days after biting an infectious human and likely remain so until death (19).

Unfortunately, many of these distributions are estimated based on fewer than 30 cases. Expansion of this pool of evidence is critical for accurate assessment of surveillance activities and modeling of ZIKV risk.

Clinical illness

ZIKV symptoms are typically nonspecific and mild. Consistent with other reports (73), symptoms reported from 31 confirmed cases on Yap

Island included maculopapular rash (90%), subjective fever (65%), arthralgia or arthritis (65%), nonpurulent conjunctivitis (55%), myalgia (48%), headache (45%), retro-orbital pain (39%), edema (19%), and vomiting (10%) (6, 10). Case reports suggest that acute symptoms of ZIKV will typically fully resolve within 1 to 2 weeks of onset (44, 60, 74–80). Deaths are rare and have primarily occurred in patients with preexisting comorbidities or who are immunocompromised (81, 82).

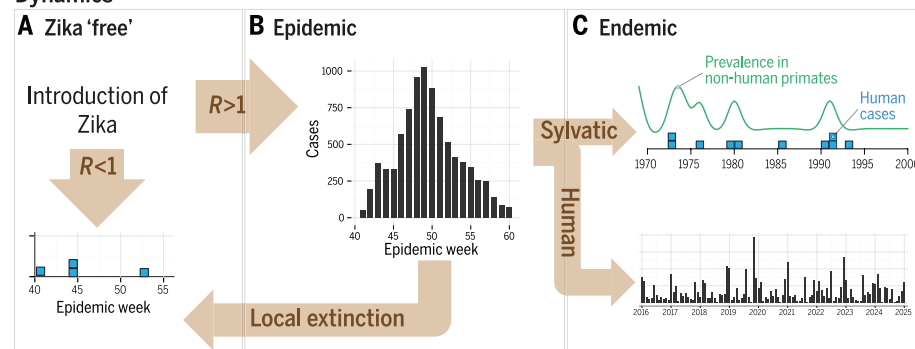
Persons infected with ZIKV may be at increased risk for severe neurologic sequelae, notably Guillain-Barré syndrome. Data from French Polynesia suggest a risk of Guillain-Barré of 24 per 100,000 ZIKV infections (5, 45), more than 10 times the annual rate in the United States (1.8 per 100,000) (83). Regardless of cause, Guillain-Barré is associated with considerable morbidity and 3 to 10% mortality (84). Guillain-Barré may be more common in symptomatic ZIKV cases; during the French Polynesia outbreak, 88% of Guillain-Barré cases reported symptoms a median of 6 days before Guillain-Barré onset (5, 45). There have been reports of other neurologic sequelae,

including meningoencephalitis (85) and acute myelitis (86), although no causal link has been established.

ZIKV in pregnancy

Much of the concern surrounding ZIKV has focused on the link between infection in pregnancy and fetal microcephaly. As of 7 May 2016, 7438 suspected microcephaly cases have been reported in Brazil since ZIKV's emergence (1326 confirmed out of 4005 investigated), versus fewer than 200 per year before the outbreak (87, 88). Quantifying the risk of microcephaly has been complicated by uncertainty in the number of ZIKV-affected pregnancies, owing to the large fraction of cases that are asymptomatic, a lack of consensus on the definition of microcephaly, and other infectious causes of microcephaly, such as cytomegalovirus and rubella (89). However, in light of multiple epidemiologic studies and the isolation of ZIKV in amniotic fluid and fetal brain tissue, the CDC confirmed a causal link between ZIKV infection during pregnancy and severe birth defects, including microcephaly in April 2016 (90). This

Dynamics



Pathogenesis

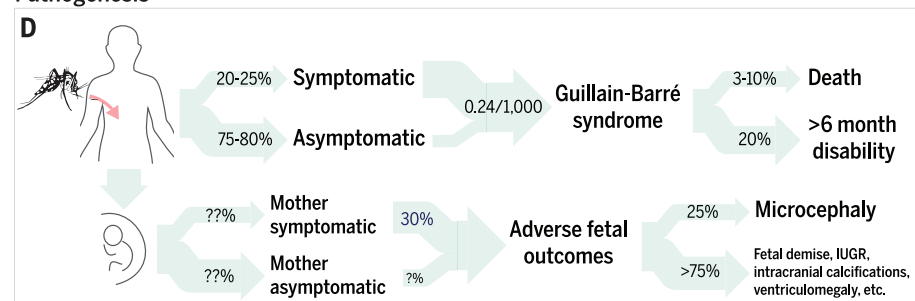


Fig. 1. Factors determining the global risk from ZIKV. (A) As long as ZIKV circulates anywhere, periodic introductions into ZIKV-free regions will occur. Whether these lead to an epidemic depends on the reproductive number R , a measure of transmission efficiency determined by local ecology and population susceptibility to ZIKV. (B) When $R > 1$, introductions can result in major epidemics, after which the virus may go locally extinct or become endemic. (C) ZIKV could be maintained endemically either in local nonhuman primates (the sylvatic cycle) or through ongoing human transmission. (D) Most ZIKV infections (75 to 80%) are asymptomatic, and those with symptoms are likely at highest risk for rare neurological complications (6, 63, 92), particularly Guillain-Barré (45). Adverse fetal outcomes, notably microcephaly, may also be more common when the mother is symptomatic. Owing to its association with pregnancy, ZIKV's health effects depend on the fertility rate and the age distribution of infections. The age distribution mirrors the general population in ZIKV-free (A) and epidemic (B) settings but is a function of the force of infection in endemic settings (C) (4, 45). Appropriate control measures can reduce R , decreasing the probability of successful ZIKV invasion (A) and its subsequent effect [(B) and (C)] [see (116)].

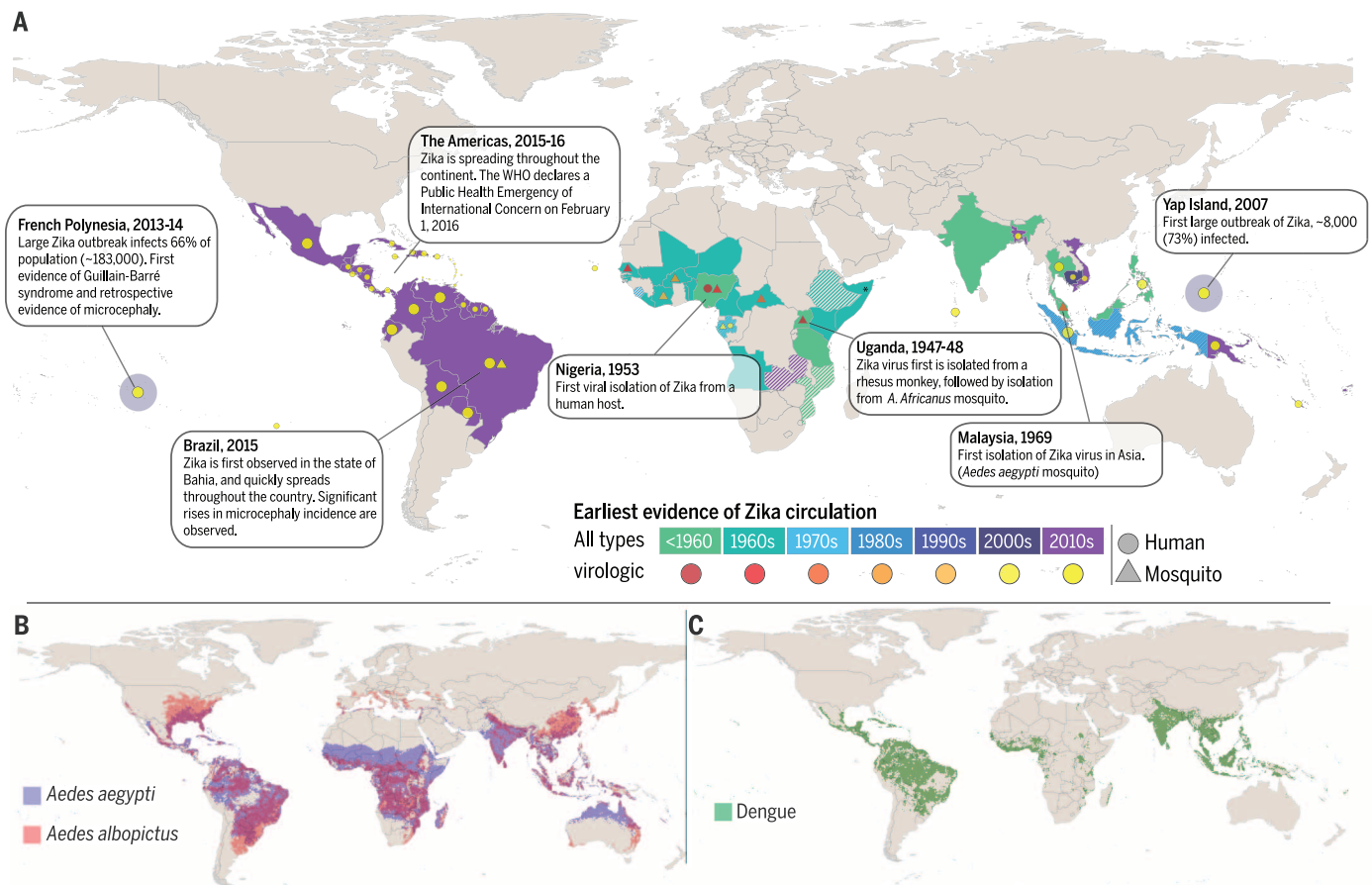


Fig. 2. Current and potential distribution of ZIKV. (A) Spread of ZIKV across the globe to date. Countries are colored by the timing of the first indication of local ZIKV transmission by serologic evidence or confirmation of human cases. Solid shading indicates clusters of confirmed cases or seropositivity to ZIKV of >10% in some subpopulations, whereas hatched colors indicate 5 to 10% seropositivity (serosurveys showing <5% seropositivity are not shown). Symbols indicate locations and timings of viral isolations from mosquitoes (triangles) and humans (circles). **(B)** Map of the global occurrence of the widely distributed

ZIKV vectors *A. aegypti* and *A. albopictus*. Adapted from (100). **(C)** Map of the occurrence of dengue, a closely related *Aedes*-transmitted flavivirus. Adapted from (103). Shaded regions correspond to areas with predicted probability of vector or dengue occurrence of >30%. *Somalia did not report the total percentage of those who were ZIKV seropositive, but there was a small percentage of subjects seropositive to ZIKV and no other flavivirus and a large percentage seropositive to two or more flaviviruses, so Somalia's data are included.

conclusion is further supported by the presence of microcephaly and other brain abnormalities in the pups of mice experimentally infected with ZIKV (97).

ZIKV symptoms in pregnant women are similar to the general population (92), but it is unknown if immunosuppression during pregnancy changes the rate at which they occur. Among those who are symptomatic, adverse fetal outcomes appear to be frequent, occurring in 29% (12 out of 42) of symptomatic ZIKV-infected pregnant women in a prospective study in Brazil (92). A second Brazilian study found that 74% (26 out of 35) of mothers of infants with microcephaly reported a rash in the first or second trimester (51). The rate of birth defects in asymptomatic pregnant women is likely lower, but not zero. For example, a Colombian study identified four microcephaly cases with virologic evidence of ZIKV infection, all of which were born to women who did not report symptoms of ZIKV (54). Modeling studies suggest that the overall risk of ZIKV-associated microcephaly in the first trimester is around 1 per 100, regardless of symptoms, and low to negligible thereafter (46, 93).

Although microcephaly was the first fetal abnormality to be recognized, there is increasing evidence that ZIKV may be responsible for other fetal sequelae, such as intracranial calcifications, ventriculomegaly, ocular impairment, brainstem hypoplasia, intrauterine growth restriction (IUGR), and fetal demise (92, 94). Placental pathology has also been reported. Although microcephaly is detectable at birth, other findings may require additional, less routine procedures, such as imaging or autopsy, and thus may be underreported. Brasil *et al.* found that only one in four fetuses with abnormalities in ZIKV-infected women met the criteria for microcephaly (92), indicating that the total number of ZIKV-affected pregnancies may be four times the number of reported microcephaly cases.

Beyond an association with symptoms, it is unclear what factors increase the risk of adverse pregnancy outcomes after maternal ZIKV infection. For other infections that cause fetal abnormalities, risk is often associated with gestational age at infection. For instance, the risk of birth defects from cytomegalovirus and rubella is high-

est if infection occurs in the first or early in the second trimester (89). Epidemiologic evidence suggests a similar association with first-trimester ZIKV infection (46, 95). In a prospective study of 88 women, microcephaly and brain abnormalities occurred only in first- and second-trimester infections (92). However, 8 of 12 cases of fetal abnormalities overall occurred in second- and third-trimester infections, and women infected as late as 35 weeks experienced fetal death, IUGR, or anhydramnios [although these outcomes commonly occur in the absence of ZIKV; e.g., in Brazil, 11 fetal deaths occur per 1000 births (96), and IUGR rates range from 5 to 7% in developed countries (97)]. A recent Colombian study suggests little to no risk from infection in the third trimester; among 616 Colombian women with clinical symptoms of ZIKV during the third trimester, none gave birth to infants with microcephaly or other brain abnormalities (7% were still pregnant at the time of reporting) (54).

Adverse outcomes in pregnancy are the most worrisome side effects of ZIKV infection, and research into this association is progressing rapidly.

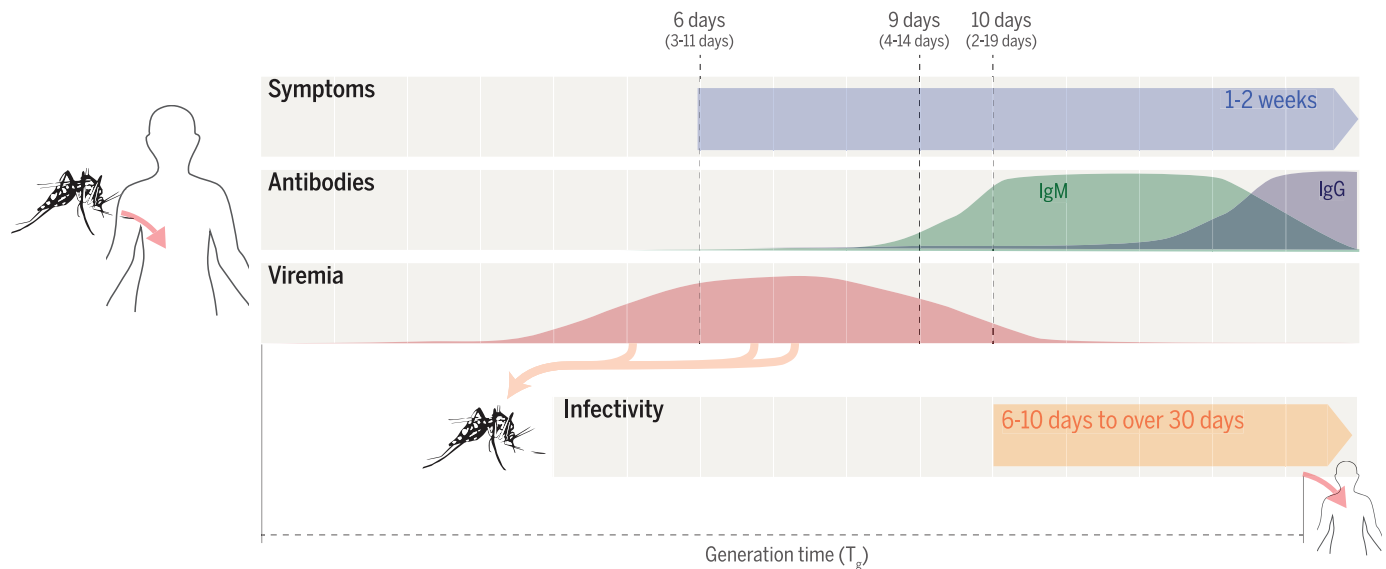


Fig. 3. Schematic of the course of human and mosquito infection. Symptoms develop, on average, 6 days (95% range, 3 to 11 days) after ZIKV infection (64). Approximately 9 days (95% range, 4 to 14 days) after infection, antibodies start increasing: The first antibodies detectable will be IgM, which will later decline as IgG antibodies increase, then persist indefinitely (the timing of the IgM/IgG switch is for illustrative purposes only and is not meant

to indicate the actual length of IgM persistence). Viremia likely starts to increase before symptoms appear, and the magnitude and length of viremia will shape the risk of infection of susceptible mosquitoes that bite this host. After an incubation period, this infected mosquito will be able to transmit infection to susceptible humans (19). The interval from the initial to the subsequent human infection is the generation time of ZIKV, T_g [for estimates, see (116)].

Still, much remains to be learned, particularly about the frequency and spectrum of ZIKV sequelae in pregnancy and how we can assess and reduce risk. ZIKV-related birth defects can have long-standing financial, social, and health effects on affected families and communities (98). Hence, the threat from ZIKV cannot purely be assessed based on immediate clinical outcomes but also must account for its lifelong effects.

The potential range and effect of ZIKV Transmissibility and potential range of ZIKV

Transmission of ZIKV in a population is a function of local ecology, the natural history of ZIKV, and the population's susceptibility to infection. The suitability of the local environment for ZIKV transmission and the effect of ZIKV's natural history are captured by the basic reproductive number R_0 , the number of secondary infections expected from a single case in a population with no preexisting immunity (e.g., French Polynesia before 2013). R_0 is a function of both disease and setting and will vary between locales based on the local environment, human behavior, vector abundance, and, potentially, interactions with other viruses. The combined effect of these factors and susceptibility will be captured by the reproductive number R , which is related to R_0 by the equation $R = R_0 \times S$, where S is the proportion of the population susceptible to ZIKV. This value, combined with the generation time (the time separating two consecutive infections in a chain of transmission), tells us the speed at which ZIKV will spread in a population. As we consider how to assess the range and effects of ZIKV, we rely both on previous experience with

ZIKV and related viruses and on an assessment of factors likely to influence R and R_0 .

The size of an outbreak after an introduction will depend on R (R_0 in a ZIKV-naïve population) (99), with small, self-limiting outbreaks becoming more likely as R approaches one, and increasing epidemics with larger R s. Hence, ZIKV can successfully spread to a new region if $R > 1$, which requires, among other factors, sufficient density of the vector population. ZIKV has been isolated from multiple *Aedes* genus mosquitoes (23–26, 38), including *A. albopictus* and *A. aegypti*, which have a large global range (Fig. 2B) (100). Although ZIKV has been occasionally isolated from or experimentally passed to other genera, including *Culex* species, there is no current evidence that they contribute substantially to its spread (22, 23, 101). It is unclear whether all areas across the range of these mosquitoes are at risk for ZIKV epidemics. Dengue, a virus that is also transmitted by *Aedes* mosquitoes, has caused epidemics throughout the Americas (Fig. 2C) but has not achieved sustained transmission in the continental United States, despite widespread vector presence (100, 102, 103). The reasons for this may include not only climate but also differences in built environments and social factors (104), all of which are likely to affect ZIKV transmission.

Several groups have attempted to map ZIKV's potential global range based on currently available data. These maps have been constructed around combinations of environmental, vector abundance, and socioeconomic factors (105–109). There is wide agreement that much of the world's tropical and subtropical regions are at risk for ZIKV spread, including major portions of the Americas, Africa, Southeast Asia, and the Indian

subcontinent, as well as many Pacific islands and Northern Australia. These maps differ notably in the extent of risk projected in the southeastern United States and inland areas of South America and Africa, with Carlson and colleagues suggesting a more limited range (107), particularly in the continental United States, than Messina *et al.* and Samy *et al.* (108, 109). These maps are important attempts to refine estimates of ZIKV's global range beyond those based solely on the distribution of dengue or *Aedes* mosquitoes but, as noted by the authors, are based on limited evidence and should be refined as we learn more about ZIKV. These analyses are, arguably, best interpreted as an assessment of the risk of initial postinvasion ZIKV epidemics, not its long-term persistence. Whether ZIKV will in fact spread throughout these areas is uncertain; similar viruses have failed to spread to or take hold in areas theoretically at risk (e.g., yellow fever in Southeast Asia) (110).

R_0 in ZIKV outbreaks in Yap Island and French Polynesia was estimated to be between 1.8 and 5.8 (111–113), corresponding to 73.2 to 99.9% of the at-risk population becoming infected in an uncontrolled outbreak, based on classic epidemic theory (4) [although the true relationship between R_0 and final attack rates for ZIKV will be somewhat more complex (99)]. Serosurveys in French Polynesia suggest that 66% of the population was infected (46), which is somewhat lower but not inconsistent with these projections. Preliminary estimates of R_0 from Colombia vary by location and range from 1.4 to 6.6 (114, 115). These are similar to R_0 estimates presented by Ferguson *et al.* for 13 countries in the Americas (116) and recent estimates of R_0 for Rio de Janeiro (117).

These values are consistent with R_0 estimates for dengue in similar settings. Of note, all of these are from settings with recently observed endogenous transmission of ZIKV, and R_0 will vary widely across settings and is likely to be far lower near the limits and outside of ZIKV's range.

ZIKV's potential for endemic circulation

After the initial, postinvasion epidemic of ZIKV, the virus may either go extinct locally or be maintained through endemic human spread or sylvatic transmission (Fig. 1). Early age-stratified serosurveys in Africa and Asia offer some insight into past transmission patterns of ZIKV in these regions and ZIKV's past dynamics (Fig. 4). Serosurveys in Nigeria, the Central African Republic, and Malaysia are consistent with ongoing ZIKV transmission, common spillover infections from a sylvatic reservoir, or frequent reintroductions from other regions over multiple decades (13, 16, 118). However, these results must be interpreted with caution owing to cross-reactivity with other flaviviruses in serologic tests (22). Up-to-date, age-stratified serosurveys, broadly

covering regions where ZIKV has previously been detected, would tell us much about the virus's ability to persist.

More recent evidence of sustained transmission comes from Thailand, where seven samples collected in independent outbreak investigations tested positive for ZIKV infection (43). The broad geographic spread of these cases is consistent with endemic transmission throughout Thailand. Furthermore, occasional but consistent serologic and virologic evidence of ZIKV transmission in humans and mosquitoes from across Africa, India, and Southeast Asia spanning more than 60 years suggests that ZIKV has been persistently present throughout these regions (22) (Fig. 1A). Phylogenetic evidence further supports this supposition, because the African and Asian lineages divided in the 1940s and remain distinct up until the present day (22, 26) (Fig. 5).

The evidence supports ZIKV's ability to persist regionally, but it is unclear whether the human population alone can maintain ZIKV endemically. After an initial postinvasion epidemic, the time until there is a risk of additional epidemics will

be driven by the replenishment of susceptibles through births and waning immunity [the latter seems unlikely based on evidence that other flaviviruses provide lifelong immunity to the infecting strain (22)]. For ZIKV to persist in the human population over this period, the population must be large enough to support low levels of transmission between epidemics (4).

However, all countries with evidence of persistent ZIKV transmission have a plausible sylvatic cycle. Patterns of ZIKV isolations in a study of samples from multiple hosts in Senegal spanning 50 years support episodic transmission across species (9); phylogenetic evidence indicates ZIKV passes frequently between nonhuman primates and humans in Africa (26); and numerous studies in Africa and Asia show serologic evidence for ZIKV infection in nonhuman primates (1, 18, 22, 33, 119). Some areas, where there has been serological evidence of long periods of consistent risk of ZIKV infection, are near areas where serological evidence suggests that human populations are largely ZIKV free (e.g., Nigeria versus Kenya) (120, 121)—a pattern more consistent with spillover infections from a sylvatic reservoir than of endemic transmission in humans.

In light of this evidence, it is plausible that the persistence of ZIKV in Africa and Asia may depend on the presence of a sustainable sylvatic cycle. However, it is unclear if the primate population in the Americas could support sylvatic transmission (122) or if such a cycle is necessary for ZIKV to remain endemic. Nonhuman primates are present throughout South and Central America, and ZIKV has recently been isolated from two species in the Ceará State of Brazil (123), suggesting at least the possibility for sustained sylvatic transmission in the region. Further characterization of ZIKV ecology in Asia and Africa and monitoring of the developing situation in the Americas is needed to assess the long-term risk from ZIKV in newly affected regions.

Because the most severe outcomes of ZIKV infection are associated with pregnancy, the risk from endemic ZIKV will depend on the age distribution of those infected. Serosurveys indicating ongoing ZIKV circulation (Fig. 4, A to C) support average ages of infection of 17 (Nigeria, 1952), 29 (Central African Republic, 1979) and 30 years (Malaysia 1953 to 1954) (13, 16, 118). Likewise, R_0 estimates from the literature are consistent with average ages of infection ranging between 10 and 38 years in the setting of endemic human-to-human transmission (although human-to-human transmission should not be necessarily assumed in the settings covered in Fig. 4, A to C). These ages suggest that in endemic settings, risk of ZIKV infection may be considerable during childbearing years. Importantly, this information could potentially be used to estimate the expected rate of microcephaly and other birth defects in regions where ZIKV becomes endemic.

Why has ZIKV invaded the Americas now?

Little is known about ZIKV's introduction into the Americas. Phylogenetic analyses indicate

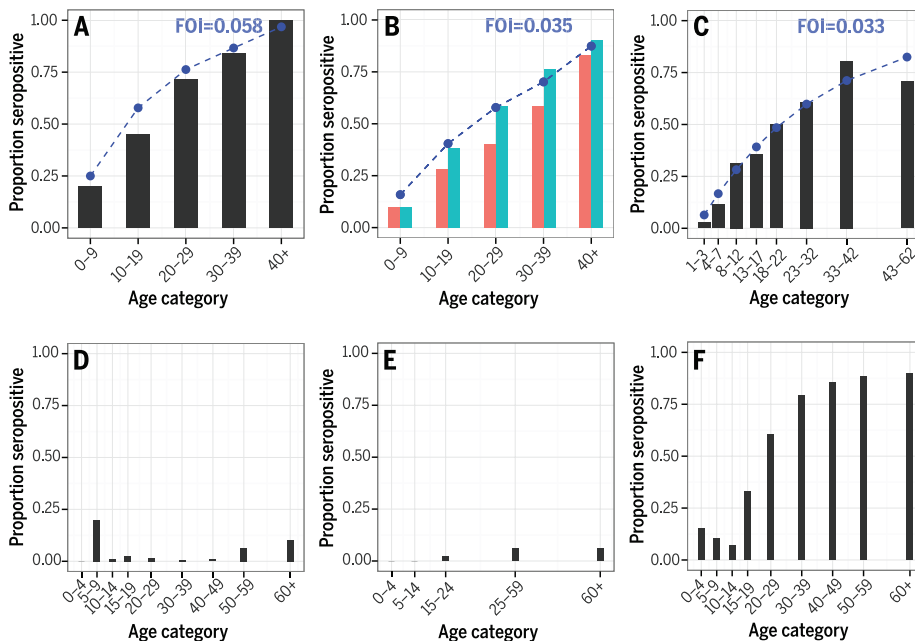


Fig. 4. Age-stratified serosurveys provide important clues to local ZIKV epidemiology. Results must be interpreted with caution because of the possibility of cross-reactivity with other flavivirus antibodies. (A to C) Ongoing ZIKV transmission, whether from endemic human transmission or a constant risk of zoonotic infection, manifests as a smooth increase in the proportion of the population seropositive with increasing age. This pattern is also consistent with frequent reintroductions leading to periodic outbreaks. If we assume that the risk of ZIKV infection is constant over a lifetime, we can estimate the force of infection (FOI), the proportion of the susceptible population infected each year. Serosurvey results consistent with ongoing transmission include (A) Uburu, Nigeria, 1952 (13); (B) Central African Republic, 1979 (pink, female; cyan, male) (118); and (C) Malaysia, 1953 to 1954 (16). Blue dashed lines and text represent the expected distribution from the estimated FOI. (D and E) In areas without substantial ZIKV transmission, there will be very low levels of seropositivity across age groups and no clear age pattern. Some individuals may still be seropositive due to cross-reactivity in serological assays, infection of travelers, and limited imported cases. Examples include (D) Central Nyanza, Kenya, 1966 to 1968 (121) and (E) Mid-Western Region, Nigeria, 1966 to 1967 (120). (F) Substantial shifts in seropositivity between age groups inconsistent with ongoing transmission suggest past epidemics—e.g., results from a 1966 to 1968 serosurvey in the Malindi district of Kenya are consistent with one or more epidemics of ZIKV occurring 15 to 30 years previously (121). Similar patterns could also occur due to differences in infection risk by age or a sharp reduction in transmission intensity at some point in the past.

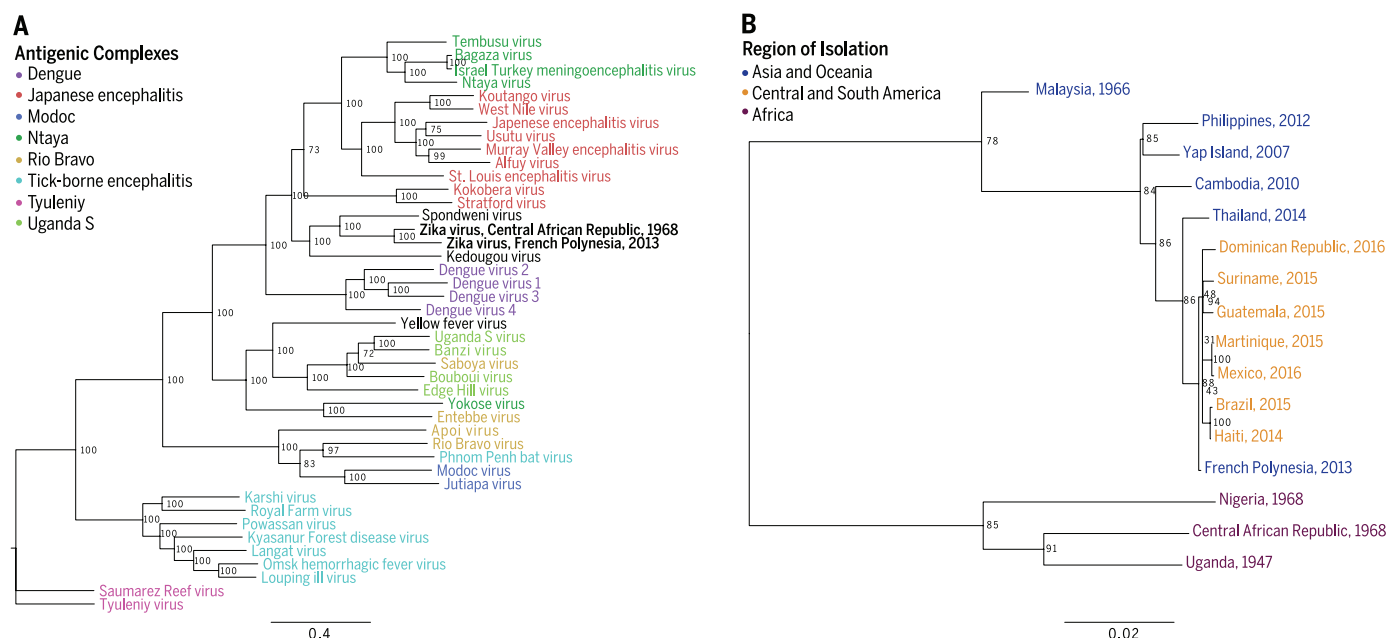


Fig. 5. ZIKV phylogenetics. (A) Maximum likelihood tree of phylogenetic relationships between 43 flaviviruses (numbers indicate support from 1000 ultrafast bootstrap replicates), with antigenic clusters from Calisher *et al.* indicated by color (162). (B) The phylogenetic relationship between ZIKV strains isolated from throughout the globe. Whole-genome nucleotide sequences were aligned using Clustal Omega (163), and trees were constructed using IQ-TREE (164) under a GTR+G+I evolutionary model.

that a virus descended from the French Polynesian ZIKV strain entered Brazil between May and December 2013 (52). Although there has been speculation about introduction during specific sporting events (52, 124), Brazil has more than 6 million visitors per year, providing numerous opportunities for ZIKV introduction. Regardless of how and when ZIKV entered the Americas, the reasons for the size and severity of this outbreak are unclear.

The unprecedented size and effect of the ZIKV epidemic in the Americas may be the natural result of a random introduction into a large population without preexisting immunity. Like the Americas, the populations of Yap Island and French Polynesia were fully susceptible when ZIKV was introduced, and both had large outbreaks infecting more than 65% of their populations (6, 45). However, on these small islands the absolute number of adverse outcomes may have been too low to be noticed initially. Likewise, it is possible that small ZIKV epidemics, and even invasion into Southeast Asia in the mid-1900s, resulted in effects that were unnoticed against the backdrop of other infectious diseases, particularly because small population sizes (compared to Brazil) mean that excess microcephaly cases would likely be in the hundreds (or less) in any given country. Endemic transmission would be even less likely to be noticed, because yearly attack rates would be a tenth again lower (Fig. 4) (116). Still, given the magnitude and severity of the outbreak in the Americas, it seems implausible that, if such outbreaks were occurring, none were observed for over 60 years. Hypothesized changes in the biological and ecological drivers of ZIKV transmission must be carefully

assessed, because they will influence how we quantify the risk from ZIKV globally.

Warmer temperatures and rainfall resulting from the 2015–2016 El Niño may have facilitated ZIKV transmission throughout the region (125) and increased the geographic range of *Aedes* mosquitoes. Warmer temperatures have been associated with more efficient transmission of related flaviviruses (126) and greater production of adult mosquitoes (127, 128). El Niño-associated periods of flooding (which increases mosquito breeding sites) and of droughts (which can increase human-mosquito interactions) may facilitate ZIKV transmission (129, 130). However, it should not be assumed that increased temperature or rainfall will universally promote ZIKV transmission, because climatic changes have complex repercussions across food webs (from plant growth to bird behavior) and the thermal effects on the virus itself are likely to be nonlinear (131). Over a longer time scale, development and urbanization has led to a proliferation of *A. aegypti* and *A. albopictus* in densely populated areas, which may have facilitated the rise of dengue in the region and may also have provided conditions that favored ZIKV spread (132).

There is some possibility that immunological interactions with other flaviviruses may be facilitating the spread or pathogenesis of ZIKV in the Americas. In dengue, preexisting antibodies to one serotype are hypothesized to enhance subsequent infections with another serotype through a mechanism known as antibody-dependent enhancement (ADE) (133). ADE may result in increased susceptibility to infection, the likelihood of developing severe disease, and the chances of transmission (134, 135). Evidence from some in

vitro experiments and epidemiological studies show both protective and enhancing effects between immunity to Japanese encephalitis and dengue (136, 137), and several in vitro studies have shown enhancement of ZIKV replication in the presence of antibodies to other flaviviruses (138, 139). Dengue has circulated throughout much of Central and South America since it reemerged 30 years ago; hence, it is possible that such interactions are contributing to the current outbreak of severe disease. However, this would raise questions as to why similar interactions have not been seen in the dengue endemic regions of Southeast Asia that also show evidence of ZIKV circulation. Studies that measure preexisting dengue and ZIKV antibodies and track clinical outcomes may help illuminate the issue.

The severity of outcomes in recent outbreaks, compared with past observations of mild disease, has led some to hypothesize that the virus has mutated to be more pathogenic (140). Recent evidence suggests distinct codon preferences between African and Asian ZIKV lineages, although adaptive genetic changes may have an effect on viral replication and titers (141), whereas the genetic diversity of viruses isolated in ZIKV-associated microcephaly cases suggest that recent mutations may not be involved (142). Epidemiologic and laboratory studies are needed to determine whether these changes have had a substantive effect on viral pathogenesis. Until the effect of ZIKV evolution is better understood, we should be careful to balance the need to learn from previous research with the possibility that the virus has fundamentally changed.

Human genetics is known to have a profound effect on the pathogenesis of many infectious

diseases (143), and there is some indication that the same could be true for flaviviruses (144, 145). While there is evidence of ancient intermixing between Polynesian and American populations (146), there are no indications of a link between ancestry and severe outcomes from ZIKV at this point. Likewise, genetic variation in *A. aegypti* is known to affect vector competence to transmit flaviviruses (147); hence, it is possible that changes in the makeup of the vector population also influence ZIKV transmission and account for regional differences in ZIKV effects.

Challenges and research priorities for responding to the ZIKV threat Surveillance and clinical outcomes

The key challenge in ZIKV surveillance is the proportion of cases that remain asymptomatic and the nonspecificity of ZIKV symptoms (148). Dengue and chikungunya are also transmitted by *Aedes* mosquitoes, cocirculate with ZIKV, and can have a similar presentation, further complicating surveillance efforts.

Laboratory testing is needed to confirm ZIKV infection. Molecular (reverse transcription polymerase chain reaction) techniques can be used to detect ZIKV in serum, saliva, and urine (67, 149). However, there are frequent cases in which testing of different fluids gives discrepant results, and additional studies are needed to assess diagnostic accuracy (67). The timing of sample collection is crucial; viral RNA is only detectable in serum for 3 to 5 days after symptom onset (~10 days after infection) but may persist longer in other fluids (59, 64, 66).

A highly specific, easily administered antibody test would be a boon to surveillance and patient care. Such a test could be used to estimate underlying ZIKV incidence and thus rates of severe outcomes, confirm infection in studies of ZIKV pathogenesis, and test for immunity to ZIKV early in pregnancy so women can know whether they are at risk. However, serological testing is complicated by potential cross-reactivity with other flaviviruses (22). Newer enzyme-linked immunosorbent assay (ELISA) tests show promise, such as an IgG-ELISA test used in French Polynesia that, despite endemic dengue circulation, found <1% ZIKV seropositivity in blood donors before the outbreak (150).

To assess the risk and determinants of ZIKV-related clinical outcomes, we need studies aimed at measuring the underlying incidence of ZIKV infection, regardless of clinical presentation (e.g., serosurveys), the spectrum of illness and risk factors for severe outcomes (e.g., cohort and case-control studies), and the effect of ZIKV over longer time scales, including the length of immunity.

Ecology and evolution

There has been a high level of global concern surrounding the threat from ZIKV. One reason the concern is so great is that we are unable to accurately assess the global threat from the virus, and differing lines of evidence point to conflicting conclusions. For instance, the range of *Aedes* mosquitoes and ecological analyses would

suggest that much of the continental United States is at risk from ZIKV, whereas recent experience with dengue and chikungunya would suggest that ZIKV is unlikely to persist in this region. To assess the epidemiologic and ecologic factors that drive global risk, there is a need for studies that more accurately assess where ZIKV circulation persists over long periods (e.g., global age-stratified serosurveys) and the ecological determinants of persistence (e.g., reservoirs, critical population size, and vector competence), as well as studies characterizing interactions between ZIKV and other flaviviruses. Across both clinical and ecological studies, it is important to evaluate the effect of host, viral, and mosquito genetics.

Interventions and control

A ZIKV vaccine may be the best way to protect at-risk populations over the long term. Vaccine development has been prioritized by the WHO and other public health agencies, and there are at least 18 active manufacturers and research institutions pursuing early stages of ZIKV vaccine development (151). However, phase 1 clinical studies are not expected to begin until the end of 2016 (151); hence, a vaccine is unlikely to become available in time to change the course of the current outbreak in the Americas.

Without a vaccine or antiviral drugs, the tools at our disposal for reducing ZIKV incidence are based on vector control and limiting ZIKV exposure. We have little direct evidence of the effectiveness of these approaches in controlling ZIKV transmission, but there are decades of experience in controlling dengue and other flaviviruses (152–154). Effective vector control is possible: Gorgas virtually eliminated yellow fever from Havana and the Panama Canal region in the early 1900s using crude and draconian methods of vector control (155). Intensive vector control in the 1950s and 1960s, including mass DDT spraying, successfully eliminated *A. aegypti* from 18 countries in the Americas, substantially reducing dengue incidence (154, 156, 157). Later, Singapore and Cuba implemented successful vector-control programs lasting decades (154, 158, 159). However, all of these efforts ultimately proved to be unsustainable, and *A. aegypti* and dengue reemerged after their discontinuation (154, 158, 159). Nevertheless, there could be benefits from even short-term elimination, but research is needed to identify sustainable policies that can protect areas from ZIKV and/or other *Aedes*-borne diseases in the long term.

There is limited evidence for the effectiveness of measures aimed at reducing individual exposure to mosquitoes for dengue control. A meta-analysis suggests that use of screens in houses reduces the odds of dengue incidence by 78%, as does combined community environmental management and use of water-container covers (152). Other interventions—such as indoor residual spraying, repellents, bed nets, and traps—showed no statistically significant effect or a negative effect (insecticide aerosols) (152). However, these results are predominantly based on observational studies, limiting the strength of the evidence they provide.

Topical insect repellents and other personal protective measures do reduce mosquito biting (160) and should decrease the risk of ZIKV infection. Some randomized trials have assessed the effect of interventions on mosquito populations with inconsistent results (152, 161), and there have been no well-designed trials assessing the effect of the common, WHO-recommended practice of space spraying or fogging to control dengue transmission (152). Well-designed experimental studies with end points of transmission and disease in humans are needed to better evaluate the effectiveness of interventions aimed at vector control and personal risk reduction.

Conclusion

The rise of ZIKV after its long persistence as a disease of apparently little importance highlights how little we truly understand about the global spread of flaviviruses and other vector-borne diseases. Over the past decades, dengue, chikungunya, West Nile virus, and now ZIKV have emerged or reemerged throughout the globe (2, 145). However, why these viruses have expanded their range, while others (e.g., yellow fever) have failed to invade areas potentially ripe for their spread, remains a mystery. New analytic and molecular tools have greatly expanded our ability to forecast risk and track the spread of these viruses, but a deep understanding of what makes one virus a global threat while another is not remains elusive. Although the important role of random chance and the continuing evolution of viral species may make precise forecasting of emerging pandemics impossible, we can continue to improve the speed with which we assess and respond to emerging threats.

The evidence highlighted in this review is both encouraging and disheartening. On the one hand, the speed with which the global community has collected and disseminated clinical, epidemiologic, and laboratory information on ZIKV after identification of the threat is impressive. But the development of therapeutics and diagnostics is hampered by our ignorance, despite knowing of ZIKV's existence for more than half a century. Consequently, we have been able to do little to contain the virus's rapid spread across the Americas. New threats from infectious diseases may emerge from unexpected places, and we need strategies in place that we can roll out to rapidly gain an understanding of the transmission, pathogenesis, and control of previously little-known pathogens to protect global public health.

REFERENCES AND NOTES

1. G. W. A. Dick, S. F. Kitchen, A. J. Haddow, Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 509–520 (1952). doi: 10.1016/0035-9203(52)90042-4; pmid: 12995440
2. A. S. Fauci, D. M. Morens, Zika virus in the Americas: Yet another arbovirus threat. *N. Engl. J. Med.* **374**, 601–604 (2016). doi: 10.1056/NEJMp1600297; pmid: 26761185
3. WHO, WHO Director-General summarizes the outcome of the Emergency Committee regarding clusters of microcephaly and Guillain-Barré syndrome (WHO, 2016); available at www.who.int/mediacentre/news/statements/2016/emergency-committee-zika-microcephaly/en/.

4. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, USA, 1991).
5. V.-M. Cao-Lormeau et al., Zika virus, French polynesia, South Pacific, 2013. *Emerg. Infect. Dis.* **20**, 1085–1086 (2014). doi: [10.3201/eid2011.141380](https://doi.org/10.3201/eid2011.141380); pmid: [24856001](https://pubmed.ncbi.nlm.nih.gov/24856001/)
6. M. R. Duffy et al., Zika virus outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* **360**, 2536–2543 (2009). doi: [10.1056/NEJMoa0805715](https://doi.org/10.1056/NEJMoa0805715); pmid: [19516034](https://pubmed.ncbi.nlm.nih.gov/19516034/)
7. E. Monlun et al., Surveillance of the circulation of arbovirus of medical interest in the region of eastern Senegal. *Bull. Soc. Pathol. Exot.* **86**, 21–28 (1993). doi: [10.1016/0035-9203\(89\)90352-0](https://doi.org/10.1016/0035-9203(89)90352-0); pmid: [2559514](https://pubmed.ncbi.nlm.nih.gov/2559514/)
8. J. O. Lloyd-Smith et al., Epidemic dynamics at the human-animal interface. *Science* **326**, 1362–1367 (2009). doi: [10.1126/science.1177345](https://doi.org/10.1126/science.1177345); pmid: [19965751](https://pubmed.ncbi.nlm.nih.gov/19965751/)
9. B. M. Althouse et al., Impact of climate and mosquito vector abundance on sylvatic arbovirus circulation dynamics in Senegal. *Am. J. Trop. Med. Hyg.* **92**, 88–97 (2015). doi: [10.4269/ajtmh.13-0617](https://doi.org/10.4269/ajtmh.13-0617); pmid: [25404071](https://pubmed.ncbi.nlm.nih.gov/25404071/)
10. L. R. Petersen, D. J. Jamieson, A. M. Powers, M. A. Honein, Zika virus. *N. Engl. J. Med.* **374**, 1552–1563 (2016). doi: [10.1056/NEJMra1602113](https://doi.org/10.1056/NEJMra1602113); pmid: [27028561](https://pubmed.ncbi.nlm.nih.gov/27028561/)
11. R. H. Kokernot, V. M. Casaca, M. P. Weinbren, B. M. McIntosh, Survey for antibodies against arthropod-borne viruses in the sera of indigenous residents of Angola. *Trans. R. Soc. Trop. Med. Hyg.* **59**, 563–570 (1965). doi: [10.1016/0035-9203\(65\)90159-8](https://doi.org/10.1016/0035-9203(65)90159-8); pmid: [5893149](https://pubmed.ncbi.nlm.nih.gov/5893149/)
12. K. C. Smithburn, Studies on certain viruses isolated in the tropics of Africa and South America; immunological reactions as determined by cross-neutralization tests. *J. Immunol.* **68**, 441–460 (1952). pmid: [14946384](https://pubmed.ncbi.nlm.nih.gov/14946384/)
13. F. N. MacNamara, Zika virus: A report on three cases of human infection during an epidemic of jaundice in Nigeria. *Trans. R. Soc. Trop. Med. Hyg.* **48**, 139–145 (1954). doi: [10.1016/0035-9203\(54\)90006-1](https://doi.org/10.1016/0035-9203(54)90006-1); pmid: [13157159](https://pubmed.ncbi.nlm.nih.gov/13157159/)
14. G. W. Dick, Epidemiological notes on some viruses isolated in Uganda; Yellow fever, Rift Valley fever, Bwamba fever, West Nile, Mengo, Semliki forest, Bunyamwera, Ntaya, Uganda S and Zika viruses. *Trans. R. Soc. Trop. Med. Hyg.* **47**, 13–48 (1953). doi: [10.1016/0035-9203\(53\)90021-2](https://doi.org/10.1016/0035-9203(53)90021-2); pmid: [13077697](https://pubmed.ncbi.nlm.nih.gov/13077697/)
15. K. C. Smithburn, J. A. Kerr, P. B. Gatne, Neutralizing antibodies against certain viruses in the sera of residents of India. *J. Immunol.* **72**, 248–257 (1954). pmid: [13163397](https://pubmed.ncbi.nlm.nih.gov/13163397/)
16. W. L. Pond, Arthropod-borne virus antibodies in sera from residents of south-east Asia. *Trans. R. Soc. Trop. Med. Hyg.* **57**, 364–371 (1963). doi: [10.1016/0035-9203\(63\)90100-7](https://doi.org/10.1016/0035-9203(63)90100-7); pmid: [14062273](https://pubmed.ncbi.nlm.nih.gov/14062273/)
17. K. C. Smithburn, Neutralizing antibodies against arthropod-borne viruses in the sera of long-time residents of Malaya and Borneo. *Am. J. Hyg.* **59**, 157–163 (1954). pmid: [13138582](https://pubmed.ncbi.nlm.nih.gov/13138582/)
18. G. W. A. Dick, Zika virus. II. Pathogenicity and physical properties. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 521–534 (1952). doi: [10.1016/0035-9203\(52\)90043-6](https://doi.org/10.1016/0035-9203(52)90043-6); pmid: [12995441](https://pubmed.ncbi.nlm.nih.gov/12995441/)
19. W. G. C. Bearcroft, Zika virus infection experimentally induced in a human volunteer. *Trans. R. Soc. Trop. Med. Hyg.* **50**, 442–448 (1956). doi: [10.1016/0035-9203\(56\)90090-6](https://doi.org/10.1016/0035-9203(56)90090-6); pmid: [13380987](https://pubmed.ncbi.nlm.nih.gov/13380987/)
20. D. I. Simpson, Zika virus infection in man. *Trans. R. Soc. Trop. Med. Hyg.* **58**, 335–337 (1964). doi: [10.1016/0035-9203\(64\)90201-9](https://doi.org/10.1016/0035-9203(64)90201-9); pmid: [14175744](https://pubmed.ncbi.nlm.nih.gov/14175744/)
21. M. Cornet, Y. Robin, C. Adam, M. Valade, M. A. Calvo, Transmission expérimentale comparée du virus amaril et du virus Zika chez *Aedes aegypti* L. *Cah. ORSTOM Série Entomologie Médicale et Parasitologie* **17**, 47–53 (1979).
22. D. Musso, D. J. Gubler, Zika virus. *Clin. Microbiol. Rev.* **29**, 487–524 (2016). doi: [10.1128/CMR.00072-15](https://doi.org/10.1128/CMR.00072-15); pmid: [27029595](https://pubmed.ncbi.nlm.nih.gov/27029595/)
23. G. Vogel, Mosquito hunters search for Zika vectors. *Science* **352**, 1152–1153 (2016). doi: [10.1126/science.352.6290.1152](https://doi.org/10.1126/science.352.6290.1152); pmid: [27257232](https://pubmed.ncbi.nlm.nih.gov/27257232/)
24. P.-S. J. Wong, M.-Z. I. Li, C.-S. Chong, L.-C. Ng, C.-H. Tan, *Aedes* (Stegomyia) albopictus (Skuse): A potential vector of Zika virus in Singapore. *PLOS Negl. Trop. Dis.* **7**, e2348 (2013). doi: [10.1371/journal.pntd.0002348](https://doi.org/10.1371/journal.pntd.0002348); pmid: [23936579](https://pubmed.ncbi.nlm.nih.gov/23936579/)
25. G. Gard et al., Zika virus in Gabon (Central Africa)—2007: A new threat from *Aedes albopictus*? *PLOS Negl. Trop. Dis.* **8**, e2681 (2014). doi: [10.1371/journal.pntd.0002681](https://doi.org/10.1371/journal.pntd.0002681); pmid: [24516683](https://pubmed.ncbi.nlm.nih.gov/24516683/)
26. O. Faye et al., Molecular evolution of Zika virus during its emergence in the 20th century. *PLOS Negl. Trop. Dis.* **8**, e2636 (2014). doi: [10.1371/journal.pntd.0002636](https://doi.org/10.1371/journal.pntd.0002636); pmid: [24421913](https://pubmed.ncbi.nlm.nih.gov/24421913/)
27. E. T. Bowen et al., Large scale irrigation and arbovirus epidemiology, Kano Plain, Kenya. II. Preliminary serological survey. *Trans. R. Soc. Trop. Med. Hyg.* **67**, 702–709 (1973). doi: [10.1016/0035-9203\(73\)90041-2](https://doi.org/10.1016/0035-9203(73)90041-2); pmid: [4779116](https://pubmed.ncbi.nlm.nih.gov/4779116/)
28. F. Rodhain et al., Arbovirus infections and viral haemorrhagic fevers in Uganda: A serological survey in Karamoja district, 1984. *Trans. R. Soc. Trop. Med. Hyg.* **83**, 851–854 (1989). doi: [10.1016/0035-9203\(89\)90352-0](https://doi.org/10.1016/0035-9203(89)90352-0); pmid: [2559514](https://pubmed.ncbi.nlm.nih.gov/2559514/)
29. A. H. Fagbami, Zika virus infections in Nigeria: Virological and seroepidemiological investigations in Oyo State. *J. Hyg.* **83**, 213–219 (1979). doi: [10.1017/S0022172400025997](https://doi.org/10.1017/S0022172400025997); pmid: [489960](https://pubmed.ncbi.nlm.nih.gov/489960/)
30. M. A. Darwish, H. Hoogstraal, T. J. Roberts, I. P. Ahmed, F. Omar, A sero-epidemiological survey for certain arboviruses (Togaviridae) in Pakistan. *Trans. R. Soc. Trop. Med. Hyg.* **77**, 442–445 (1983). doi: [10.1016/0035-9203\(83\)90106-2](https://doi.org/10.1016/0035-9203(83)90106-2); pmid: [6314612](https://pubmed.ncbi.nlm.nih.gov/6314612/)
31. J. G. Olson, T. G. Ksiazek, Suhandiman, Triwibowo, Zika virus, a cause of fever in Central Java, Indonesia. *Trans. R. Soc. Trop. Med. Hyg.* **75**, 389–393 (1981). doi: [10.1016/0035-9203\(81\)90100-0](https://doi.org/10.1016/0035-9203(81)90100-0); pmid: [6275577](https://pubmed.ncbi.nlm.nih.gov/6275577/)
32. J. G. Olson et al., A survey for arboviral antibodies in sera of humans and animals in Lombok, Republic of Indonesia. *Ann. Trop. Med. Parasitol.* **77**, 131–137 (1983). pmid: [6309104](https://pubmed.ncbi.nlm.nih.gov/6309104/)
33. A. W. McCrae, B. G. Kirya, Yellow fever and Zika virus epizootics and enzootics in Uganda. *Trans. R. Soc. Trop. Med. Hyg.* **76**, 552–562 (1982). doi: [10.1016/0035-9203\(82\)90161-4](https://doi.org/10.1016/0035-9203(82)90161-4); pmid: [6304948](https://pubmed.ncbi.nlm.nih.gov/6304948/)
34. V. H. Lee, D. L. Moore, Vectors of the 1969 yellow fever epidemic on the Jos Plateau, Nigeria. *Bull. World Health Organ.* **46**, 669–673 (1972). pmid: [4403105](https://pubmed.ncbi.nlm.nih.gov/4403105/)
35. A. J. Haddow, M. C. Williams, J. P. Woodall, D. I. Simpson, L. K. Goma, Twelve isolations of Zika virus from *Aedes* (Stegomyia) africanus (Theobald) taken in and above a Uganda forest. *Bull. World Health Organ.* **31**, 57–69 (1964). pmid: [14230895](https://pubmed.ncbi.nlm.nih.gov/14230895/)
36. B. M. McIntosh, C. B. Worth, R. H. Kokernot, Isolation of Semliki Forest virus from *Aedes* (Aedimorphus) argenteopunctatus (Theobald) collected in Portuguese East Africa. *Trans. R. Soc. Trop. Med. Hyg.* **55**, 192–198 (1961). doi: [10.1016/0035-9203\(61\)90025-6](https://doi.org/10.1016/0035-9203(61)90025-6); pmid: [13774007](https://pubmed.ncbi.nlm.nih.gov/13774007/)
37. M. P. Weinbren, M. C. Williams, Zika virus: Further isolations in the Zika area, and some studies on the strains isolated. *Trans. R. Soc. Trop. Med. Hyg.* **52**, 263–268 (1958). doi: [10.1016/0035-9203\(58\)90085-3](https://doi.org/10.1016/0035-9203(58)90085-3); pmid: [13556872](https://pubmed.ncbi.nlm.nih.gov/13556872/)
38. N. J. Marchette, R. Garcia, A. Rudnick, Isolation of Zika virus from *Aedes aegypti* mosquitoes in Malaysia. *Am. J. Trop. Med. Hyg.* **18**, 411–415 (1969). doi: [10.4269/ajtmh.13-0029](https://doi.org/10.4269/ajtmh.13-0029); pmid: [23878182](https://pubmed.ncbi.nlm.nih.gov/23878182/)
39. A. Fagbami, Epidemiological investigations on arbovirus infections at Igbo-Ora, Nigeria. *Trop. Geogr. Med.* **29**, 187–191 (1977). pmid: [906078](https://pubmed.ncbi.nlm.nih.gov/906078/)
40. D. L. Moore et al., Arthropod-borne viral infections of man in Nigeria, 1964–1970. *Ann. Trop. Med. Parasitol.* **69**, 49–64 (1975). doi: [10.1080/00034983.1975.11686983](https://doi.org/10.1080/00034983.1975.11686983); pmid: [1124969](https://pubmed.ncbi.nlm.nih.gov/1124969/)
41. R. S. Lanciotti et al., Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* **14**, 1232–1239 (2008). doi: [10.3201/eid1408.080287](https://doi.org/10.3201/eid1408.080287); pmid: [18680646](https://pubmed.ncbi.nlm.nih.gov/18680646/)
42. J. C. Kwong, J. D. Druce, K. Leder, Zika virus infection acquired during brief travel to Indonesia. *Am. J. Trop. Med. Hyg.* **89**, 516–517 (2013). doi: [10.4269/ajtmh.13-0029](https://doi.org/10.4269/ajtmh.13-0029); pmid: [23878182](https://pubmed.ncbi.nlm.nih.gov/23878182/)
43. R. Buathong et al., Detection of Zika virus infection in Thailand, 2012–2014. *Am. J. Trop. Med. Hyg.* **93**, 380–383 (2015). doi: [10.4269/ajtmh.15-0022](https://doi.org/10.4269/ajtmh.15-0022); pmid: [26101272](https://pubmed.ncbi.nlm.nih.gov/26101272/)
44. K. Fonseca et al., First case of Zika virus infection in a returning Canadian traveler. *Am. J. Trop. Med. Hyg.* **91**, 1035–1038 (2014). doi: [10.4269/ajtmh.14-0151](https://doi.org/10.4269/ajtmh.14-0151); pmid: [25294619](https://pubmed.ncbi.nlm.nih.gov/25294619/)
45. V.-M. Cao-Lormeau et al., Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: A case-control study. *Lancet* **387**, 1531–1539 (2016). doi: [10.1016/S0140-6736\(16\)00562-6](https://doi.org/10.1016/S0140-6736(16)00562-6); pmid: [26948433](https://pubmed.ncbi.nlm.nih.gov/26948433/)
46. S. Cauchemez et al., Association between Zika virus and microcephaly in French Polynesia, 2013–15: A retrospective study. *Lancet* **387**, 2125–2132 (2016). doi: [10.1016/S0140-6736\(16\)00651-6](https://doi.org/10.1016/S0140-6736(16)00651-6); pmid: [26993883](https://pubmed.ncbi.nlm.nih.gov/26993883/)
47. A. Roth et al., Concurrent outbreaks of dengue, chikungunya and Zika virus infections - an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012–2014. *Euro Surveill.* **19**, 20929 (2014). doi: [10.2807/1560-7917.ES2014.19.41.20929](https://doi.org/10.2807/1560-7917.ES2014.19.41.20929); pmid: [25345518](https://pubmed.ncbi.nlm.nih.gov/25345518/)
48. G. S. Campos, A. C. Bandeira, S. I. Sardi, Zika virus outbreak, Bahia, Brazil. *Emerg. Infect. Dis.* **21**, 1885–1886 (2015). doi: [10.3201/eid2110.150847](https://doi.org/10.3201/eid2110.150847); pmid: [26401719](https://pubmed.ncbi.nlm.nih.gov/26401719/)
49. J. Lednicky et al., Zika virus outbreak in Haiti in 2014: Molecular and clinical data. *PLOS Negl. Trop. Dis.* **10**, e0004687 (2016). doi: [10.1371/journal.pntd.0004687](https://doi.org/10.1371/journal.pntd.0004687); pmid: [27111294](https://pubmed.ncbi.nlm.nih.gov/27111294/)
50. World Health Organization, Zika virus outbreaks in the Americas. *Wkly. Epidemiol. Rec.* **90**, 609–610 (2015). pmid: [26552108](https://pubmed.ncbi.nlm.nih.gov/26552108/)
51. L. Schuler-Faccini et al., Brazilian Medical Genetics Society—Zika Embryopathy Task Force, Possible association between Zika virus infection and microcephaly - Brazil, 2015. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 59–62 (2016). doi: [10.15585/mmwr.mm6503e2](https://doi.org/10.15585/mmwr.mm6503e2); pmid: [26820244](https://pubmed.ncbi.nlm.nih.gov/26820244/)
52. N. R. Faria et al., Zika virus in the Americas: Early epidemiological and genetic findings. *Science* **352**, 345–349 (2016). pmid: [27013429](https://pubmed.ncbi.nlm.nih.gov/27013429/)
53. CDC, All Countries and Territories with Active Zika Virus Transmission; available at www.cdc.gov/zika/geo/active-countries.html.
54. O. Pacheco et al., Zika virus disease in Colombia: Preliminary report. *N. Engl. J. Med.* **NEJMoa1604037** (2016). doi: [10.1056/NEJMoa1604037](https://doi.org/10.1056/NEJMoa1604037); pmid: [27305043](https://pubmed.ncbi.nlm.nih.gov/27305043/)
55. N. Casey, M. E. Díaz, Colombia reports first cases of microcephaly linked to Zika virus. *New York Times* (14 April, 2016); available at www.nytimes.com/2016/04/15/world/americas/columbia-reports-first-cases-of-microcephaly-linked-to-zika-virus.html?smid=pl-share.
56. Boletín Epidemiológico Semanal, Semana epidemiológica número 12 de 2016; available at www.ins.gov.co/boletin-epidemiologico/Boletn%20Epidemiologico/2016%20BoletnC3%ADn%20epidemiologicoC3%B3gico%20semana%2012.pdf.
57. J. F. W. Chan, G. K. Y. Choi, C. C. Y. Yip, V. C. C. Cheng, K.-Y. Yuen, Zika fever and congenital Zika syndrome: An unexpected emerging arboviral disease. *J. Infect.* **72**, 507–524 (2016). doi: [10.1016/j.jinf.2016.02.011](https://doi.org/10.1016/j.jinf.2016.02.011); pmid: [26940504](https://pubmed.ncbi.nlm.nih.gov/26940504/)
58. D. T. Decker et al., Male-to-male sexual transmission of Zika virus - Texas, January 2016. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 372–374 (2016). doi: [10.15585/mmwr.mm6514a3](https://doi.org/10.15585/mmwr.mm6514a3); pmid: [27078057](https://pubmed.ncbi.nlm.nih.gov/27078057/)
59. G. Venturi et al., An autochthonous case of Zika due to possible sexual transmission, Florence, Italy, 2014. *Euro Surveill.* **21**, 30148 (2016). doi: [10.2807/1560-7917.ES2016.21.8.30148](https://doi.org/10.2807/1560-7917.ES2016.21.8.30148); pmid: [26939607](https://pubmed.ncbi.nlm.nih.gov/26939607/)
60. B. D. Foy et al., Probable non-vector-borne transmission of Zika virus, Colorado, USA. *Emerg. Infect. Dis.* **17**, 880–882 (2011). doi: [10.3201/eid1705.101939](https://doi.org/10.3201/eid1705.101939); pmid: [21529401](https://pubmed.ncbi.nlm.nih.gov/21529401/)
61. M. Besnard, S. Lestere, A. Teissier, V. Cao-Lormeau, D. Musso, Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. *Euro Surveill.* **19**, 20751 (2014). doi: [10.2807/1560-7917.ES2014.19.13.20751](https://doi.org/10.2807/1560-7917.ES2014.19.13.20751); pmid: [24721538](https://pubmed.ncbi.nlm.nih.gov/24721538/)
62. R. Editorial, Brazil reports Zika infection from blood transfusions. *Reuters* (2016); available at www.reuters.com/article/us-health-zika-brazil-blood-idUSKCN0VD22N.
63. D. Musso et al., Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. *Euro Surveill.* **19**, 20761 (2014). doi: [10.2807/1560-7917.ES2014.19.14.20761](https://doi.org/10.2807/1560-7917.ES2014.19.14.20761); pmid: [24739982](https://pubmed.ncbi.nlm.nih.gov/24739982/)
64. J. Lessler et al., Times to Key Events in the Course of Zika Infection and Their Implications for Surveillance: A Systematic Review and Pooled Analysis. *bioRxiv* (2016), p. 041913.
65. A.-C. Gournat, O. O'Connor, E. Calvez, C. Goarant, M. Dupont-Rouzeyrol, Detection of Zika virus in urine. *Emerg. Infect. Dis.* **21**, 84–86 (2015). doi: [10.3201/eid2101.140894](https://doi.org/10.3201/eid2101.140894); pmid: [25530324](https://pubmed.ncbi.nlm.nih.gov/25530324/)
66. B. Rozé et al., Zika virus detection in urine from patients with Guillain-Barré syndrome on Martinique, January 2016. *Euro Surveill.* **21**, 2016 (2016). doi: [10.2807/1560-7917.ES2016.21.9.30154](https://doi.org/10.2807/1560-7917.ES2016.21.9.30154); pmid: [26967758](https://pubmed.ncbi.nlm.nih.gov/26967758/)
67. D. Musso et al., Detection of Zika virus in saliva. *J. Clin. Virol.* **68**, 53–55 (2015). doi: [10.1016/j.jcv.2015.04.021](https://doi.org/10.1016/j.jcv.2015.04.021); pmid: [26071336](https://pubmed.ncbi.nlm.nih.gov/26071336/)
68. D. Musso et al., Potential sexual transmission of Zika virus. *Emerg. Infect. Dis.* **21**, 359–361 (2015). doi: [10.3201/eid2102.141363](https://doi.org/10.3201/eid2102.141363); pmid: [25625872](https://pubmed.ncbi.nlm.nih.gov/25625872/)

69. G. Calvet *et al.*, Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: A case study. *Lancet Infect. Dis.* **16**, 653–660 (2016). doi: [10.1016/S1473-3099\(16\)00095-5](https://doi.org/10.1016/S1473-3099(16)00095-5); pmid: 26897108
70. R. B. Martinez *et al.*, Notes from the field: Evidence of Zika virus infection in brain and placental tissues from two congenitally infected newborns and two fetal losses—Brazil, 2015. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 159–160 (2016). doi: [10.15585/mmwr.mm6506e1](https://doi.org/10.15585/mmwr.mm6506e1); pmid: 26890059
71. R. W. Driggers *et al.*, Zika virus infection with prolonged maternal viremia and fetal brain abnormalities. *N. Engl. J. Med.* **374**, 2142–2151 (2016). doi: [10.1056/NEJMoal601824](https://doi.org/10.1056/NEJMoal601824); pmid: 27028667
72. B. R. Murphy, S. S. Whitehead, Immune response to dengue virus and prospects for a vaccine. *Annu. Rev. Immunol.* **29**, 587–619 (2011). doi: [10.1146/annurev-immunol-031210-101315](https://doi.org/10.1146/annurev-immunol-031210-101315); pmid: 21219187
73. P. Brasil *et al.*, Zika virus outbreak in Rio de Janeiro, Brazil: Clinical characterization, epidemiological and virological aspects. *PLOS Negl. Trop. Dis.* **10**, e0004636 (2016). doi: [10.1371/journal.pntd.0004636](https://doi.org/10.1371/journal.pntd.0004636); pmid: 27070912
74. D. Gyurech *et al.*, False positive dengue NS1 antigen test in a traveller with an acute Zika virus infection imported into Switzerland. *Swiss Med. Wkly.* **146**, w14296 (2016). pmid: 26859285
75. L. H. Chen, Zika virus infection in a Massachusetts resident after travel to Costa Rica: A case report. *Ann. Intern. Med.* **164**, 574–576 (2016). doi: [10.7326/L16-0075](https://doi.org/10.7326/L16-0075); pmid: 26864175
76. L. Zammarchi *et al.*, Zika virus infections imported to Italy: Clinical, immunological and virological findings, and public health implications. *J. Clin. Virol.* **63**, 32–35 (2015). doi: [10.1016/j.jcv.2014.12.005](https://doi.org/10.1016/j.jcv.2014.12.005); pmid: 25600600
77. D. Tappe *et al.*, Acute Zika virus infection after travel to Malaysian Borneo, September 2014. *Emerg. Infect. Dis.* **21**, 911–913 (2015). doi: [10.3201/eid2105.141960](https://doi.org/10.3201/eid2105.141960); pmid: 25898277
78. D. J. Summers, R. W. Acosta, A. M. Acosta, Zika virus in an American recreational traveler. *J. Travel Med.* **22**, 338–340 (2015). doi: [10.1111/jtm.12208](https://doi.org/10.1111/jtm.12208); pmid: 25996909
79. L. Zammarchi *et al.*, Zika virus infection in a traveller returning to Europe from Brazil, March 2015. *Euro Surveill.* **20**, 21153 (2015). doi: [10.2807/1560-7917.ES2015.20.23.21153](https://doi.org/10.2807/1560-7917.ES2015.20.23.21153); pmid: 26084316
80. A. T. Maria *et al.*, Zika virus infections in three travellers returning from South America and the Caribbean respectively, to Montpellier, France, December 2015 to January 2016. *Euro Surveill.* **21**, 30131 (2016). doi: [10.2807/1560-7917.ES.2016.21.6.30131](https://doi.org/10.2807/1560-7917.ES.2016.21.6.30131); pmid: 26898198
81. A. Sarmiento-Ospina, H. Vásquez-Serna, C. E. Jimenez-Canizales, W. E. Villamil-Gómez, A. J. Rodríguez-Morales, Zika virus associated deaths in Colombia. *Lancet Infect. Dis.* **16**, 523–524 (2016). doi: [10.1016/S1473-3099\(16\)30006-8](https://doi.org/10.1016/S1473-3099(16)30006-8); pmid: 27068488
82. L. Arzuza-Ortega *et al.*, Fatal sickle cell disease and Zika virus infection in girl from Colombia. *Emerg. Infect. Dis.* **22**, 925–927 (2016). doi: [10.3201/eid2205.151934](https://doi.org/10.3201/eid2205.151934); pmid: 27089120
83. A. Alshekhlee, Z. Hussain, B. Sultan, B. Katirji, Guillain-Barré syndrome: Incidence and mortality rates in US hospitals. *Neurology* **70**, 1608–1613 (2008). doi: [10.1212/01.wnl.0000310983.38724.d4](https://doi.org/10.1212/01.wnl.0000310983.38724.d4); pmid: 18443311
84. P. A. van Doorn, L. Ruts, B. C. Jacobs, Clinical features, pathogenesis, and treatment of Guillain-Barré syndrome. *Lancet Neurol.* **7**, 939–950 (2008). doi: [10.1016/S1474-4422\(08\)70215-1](https://doi.org/10.1016/S1474-4422(08)70215-1); pmid: 18848313
85. G. Carteaux *et al.*, Zika virus associated with meningoencephalitis. *N. Engl. J. Med.* **374**, 1595–1596 (2016). doi: [10.1056/NEJMc1602964](https://doi.org/10.1056/NEJMc1602964); pmid: 26958738
86. S. Mécharles *et al.*, Acute myelitis due to Zika virus infection. *Lancet* **387**, 1481 (2016). doi: [10.1016/S0140-6736\(16\)00644-9](https://doi.org/10.1016/S0140-6736(16)00644-9); pmid: 26946926
87. C. G. Victora *et al.*, Microcephaly in Brazil: How to interpret reported numbers? *Lancet* **387**, 621–624 (2016). doi: [10.1016/S0140-6736\(16\)00273-7](https://doi.org/10.1016/S0140-6736(16)00273-7); pmid: 26864961
88. C. D. E. O. de Emergências Em Saúde Pública Sobre Microcefalias, “Informe Epidemiológico No 25 – Semana Epidemiológica (Se) 18/2016 (01/05 A 07/05/2016) Monitoramento Dos Casos De Microcefalia No Brasil” (Ministerio de Salud, Brasil, 2016).
89. R. M. Kliegman, B. M. D. Stanton, J. St. Geme, N. F. Schor, *Nelson Textbook of Pediatrics* (Elsevier Health Sciences, 2015).
90. S. A. Rasmussen, D. J. Jamieson, M. A. Honein, L. R. Petersen, Zika virus and birth defects: Reviewing the evidence for causality. *N. Engl. J. Med.* **374**, 1981–1987 (2016). doi: [10.1056/NEJMsrl604338](https://doi.org/10.1056/NEJMsrl604338); pmid: 27074377
91. F. R. Cugola *et al.*, The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* **534**, 267–271 (2016). pmid: 27279226
92. P. Brasil *et al.*, Zika virus infection in pregnant women in Rio de Janeiro - Preliminary report. *N. Engl. J. Med.* **NEJMoal602412** (2016). doi: [10.1056/NEJMoal602412](https://doi.org/10.1056/NEJMoal602412); pmid: 26943629
93. M. A. Johansson, L. Mier-Y-Teran-Romero, J. Reefhuis, S. M. Gilboa, S. L. Hills, Zika and the risk of microcephaly. *N. Engl. J. Med.* **375**, 1–4 (2016). doi: [10.1056/NEJMp1605367](https://doi.org/10.1056/NEJMp1605367); pmid: 2722919
94. M. de Fatima Vasco Aragao *et al.*, Clinical features and neuroimaging (CT and MRI) findings in presumed Zika virus related congenital infection and microcephaly: Retrospective case series study. *BMJ* **353**, i1901 (2016). doi: [10.1136/bmj.i1901](https://doi.org/10.1136/bmj.i1901); pmid: 27075009
95. W. Kleber de Oliveira *et al.*, Increase in reported prevalence of microcephaly in infants born to women living in areas with confirmed Zika virus transmission during the first trimester of pregnancy - Brazil, 2015. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 242–247 (2016). doi: [10.15585/mmwr.mm6509e2](https://doi.org/10.15585/mmwr.mm6509e2); pmid: 26963593
96. F. M. Barbeiro *et al.*, Fetal deaths in Brazil: A systematic review. *Rev. Saude Publica* **49**, 22 (2015). doi: [10.1590/S0034-8910.2015049005568](https://doi.org/10.1590/S0034-8910.2015049005568); pmid: 25902565
97. World Health Organization, *Maternal Anthropometry and Pregnancy Outcomes: A WHO Collaborative Study* (WHO, 1995).
98. N. J. Waitzman, P. S. Romano, R. M. Scheffler, Estimates of the economic costs of birth defects. *Inquiry* **31**, 188–205 (1994). pmid: 8021024
99. A. Perkins, A. Siraj, C. Warren Ruktanonchai, M. Kraemer, A. Tatem, Model-based projections of Zika virus infections in childbearing women in the Americas (2016); <http://biorxiv.org/content/early/2016/05/19/039610>.
100. M. U. G. Kraemer *et al.*, The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife* **4**, e08347 (2015). doi: [10.7554/eLife.08347](https://doi.org/10.7554/eLife.08347); pmid: 26126267
101. D. Diallo *et al.*, Zika virus emergence in mosquitoes in southeastern Senegal, 2011. *PLOS ONE* **9**, e109442 (2014). doi: [10.1371/journal.pone.0109442](https://doi.org/10.1371/journal.pone.0109442); pmid: 25310102
102. Centers for Disease Control and Prevention (CDC), Locally acquired dengue, Key West, Florida, 2009–2010. *MMWR Morb. Mortal. Wkly. Rep.* **59**, 577–581 (2010). pmid: 20489680
103. S. Bhatt *et al.*, The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013). doi: [10.1038/nature12060](https://doi.org/10.1038/nature12060); pmid: 23563266
104. P. Reiter *et al.*, Texas lifestyle limits transmission of dengue virus. *Emerg. Infect. Dis.* **9**, 86–89 (2003). doi: [10.3201/eid0901.020220](https://doi.org/10.3201/eid0901.020220); pmid: 12533286
105. I. I. Bogoch *et al.*, Anticipating the international spread of Zika virus from Brazil. *Lancet* **387**, 335–336 (2016). doi: [10.1016/S0140-6736\(16\)00080-5](https://doi.org/10.1016/S0140-6736(16)00080-5); pmid: 26777915
106. A. J. Monaghan *et al.*, On the seasonal occurrence and abundance of the Zika virus vector mosquito *Aedes aegypti* in the contiguous United States. *PLOS Curr.* **8**, (2016). doi: [10.1371/currents.outbreaks.50dfc7f46798675fc63e7d7da563da76](https://doi.org/10.1371/currents.outbreaks.50dfc7f46798675fc63e7d7da563da76); pmid: 27066299
107. C. Carlson, C. Colin, D. Eric, G. Wayne, An ecological assessment of the pandemic threat of Zika virus (2016); <http://biorxiv.org/content/early/2016/03/09/040386>.
108. J. P. Messina *et al.*, Mapping global environmental suitability for Zika virus. *eLife* **5**, e15272 (2016). doi: [10.7554/eLife.15272](https://doi.org/10.7554/eLife.15272); pmid: 27090089
109. A. Samy, S. M. Thomas, A. A. El Wahed, K. P. Cohoon, A. T. Peterson, Global map of Zika virus. *Mem. Inst. Oswaldo Cruz* (2016). doi: [10.1590/0074-02760160149](https://doi.org/10.1590/0074-02760160149)
110. D. J. Rogers, A. J. Wilson, S. I. Hay, A. J. Graham, The global distribution of yellow fever and dengue. *Adv. Parasitol.* **62**, 181–220 (2006). doi: [10.1016/S0065-308X\(05\)62006-4](https://doi.org/10.1016/S0065-308X(05)62006-4); pmid: 16647971
111. A. J. Kucharski *et al.*, Transmission dynamics of Zika virus in island populations: A modelling analysis of the 2013–14 French Polynesia outbreak. *PLOS Negl. Trop. Dis.* **10**, e0004726 (2016). doi: [10.1371/journal.pntd.0004726](https://doi.org/10.1371/journal.pntd.0004726); pmid: 27186984
112. S. Funk *et al.*, Comparative analysis of dengue and Zika outbreaks reveals differences by setting and virus (2016); <http://biorxiv.org/content/early/2016/04/26/043265>.
113. H. Nishiura, K. Kinoshita, K. Mizumoto, Y. Yasuda, K. Nah, Transmission potential of Zika virus infection in the South Pacific. *Int. J. Infect. Dis.* **45**, 95–97 (2016). doi: [10.1016/j.ijid.2016.02.017](https://doi.org/10.1016/j.ijid.2016.02.017); pmid: 26923081
114. D. P. Rojas *et al.*, The epidemiology and transmissibility of Zika virus in Girardot and San Andres Island, Colombia (2016); <http://biorxiv.org/content/early/2016/04/24/049957>.
115. H. Nishiura, K. Mizumoto, W. E. Villamil-Gómez, A. J. Rodríguez-Morales, Preliminary estimation of the basic reproduction number of Zika virus infection during Colombia epidemic, 2015–2016. *Travel Med. Infect. Dis.* **14**, 274–276 (2016). doi: [10.1016/j.tmaid.2016.03.016](https://doi.org/10.1016/j.tmaid.2016.03.016); pmid: 27060613
116. N. M. Ferguson *et al.*, Understanding the invasion dynamics of Zika in Latin America: Implications for policy. *Science* **353**, 353–354 (2016). doi: [10.1126/science.aag0219](https://doi.org/10.1126/science.aag0219); pmid: 27417493
117. L. Bastos *et al.*, Zika in Rio de Janeiro: Assessment of basic reproductive number and its comparison with dengue (2016); <http://biorxiv.org/content/early/2016/05/25/055475>.
118. J. F. Saluzzo, J. P. Gonzalez, J. P. Hervé, A. J. Georges, Serological survey for the prevalence of certain arboviruses in the human population of the south-east area of Central African Republic (author's transl). *Bull. Soc. Pathol. Exot. Filiales* **74**, 490–499 (1981). pmid: 6274526
119. A. M. Kilbourn *et al.*, Health evaluation of free-ranging and semi-captive orangutans (*Pongo pygmaeus pygmaeus*) in Sabah, Malaysia. *J. Wildl. Dis.* **39**, 73–83 (2003). doi: [10.7589/0090-3558.39.1.73](https://doi.org/10.7589/0090-3558.39.1.73); pmid: 12685070
120. P. Brès, Recent data from serological surveys on the prevalence of arbovirus infections in Africa, with special reference to yellow fever. *Bull. World Health Organ.* **43**, 223–267 (1970). pmid: 5312522
121. A. Geser, B. E. Henderson, S. Christensen, A multipurpose serological survey in Kenya. 2. Results of arbovirus serological tests. *Bull. World Health Organ.* **43**, 539–552 (1970). pmid: 5313066
122. B. Althouse *et al.*, Potential for Zika virus to establish a sylvatic transmission cycle in the Americas (2016); <http://biorxiv.org/content/early/2016/04/05/047175>.
123. S. Favoretto *et al.*, First detection of Zika virus in neotropical primates in Brazil: a possible new reservoir (2016); <http://biorxiv.org/content/early/2016/04/20/049395>.
124. D. Musso, Zika virus transmission from French Polynesia to Brazil. *Emerg. Infect. Dis.* **21**, 1887 (2015). doi: [10.3201/eid2110.151125](https://doi.org/10.3201/eid2110.151125); pmid: 26403318
125. Climate Prediction Center, ENSO Diagnostic Discussion; available at www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensodisc.html.
126. D. M. Watts, D. S. Burke, B. A. Harrison, R. E. Whitmore, A. Nisalak, Effect of temperature on the vector efficiency of *Aedes aegypti* for dengue 2 virus. *Am. J. Trop. Med. Hyg.* **36**, 143–152 (1987). pmid: 3812879
127. H. M. Yang, M. L. G. Macoris, K. C. Galvani, M. T. M. Andighetti, D. M. V. Wanderley, Assessing the effects of temperature on the population of *Aedes aegypti*, the vector of dengue. *Epidemiol. Infect.* **137**, 1188–1202 (2009). doi: [10.1017/S0950268809002040](https://doi.org/10.1017/S0950268809002040); pmid: 19192322
128. B. W. Alto, S. A. Juliano, Precipitation and temperature effects on populations of *Aedes albopictus* (Diptera: Culicidae): implications for range expansion. *J. Med. Entomol.* **38**, 646–656 (2001). doi: [10.1603/0022-2585.38.5.646](https://doi.org/10.1603/0022-2585.38.5.646); pmid: 11580037
129. South America Summer Forecast: El Nino to Bring Flooding Rain to Argentina, Uruguay and Southeast Brazil. *AccuWeather*; available at www.accuweather.com/en/weather-news/south-america-summer-forecast-2015-2016/53158136.
130. R. J. Pontes, J. Freeman, J. W. Oliveira-Lima, J. C. Hodgson, A. Spielman, Vector densities that potentiate dengue outbreaks in a Brazilian city. *Am. J. Trop. Med. Hyg.* **62**, 378–383 (2000). pmid: 11037781
131. C. W. Morin, A. C. Comrie, K. Ernst, Climate and dengue transmission: Evidence and implications. *Environ. Health Perspect.* **121**, 1264–1272 (2013). pmid: 24058050
132. C. M. Rios-Velázquez *et al.*, Distribution of dengue vectors in neighborhoods with different urbanization types of Manaus, state of Amazonas, Brazil. *Mem. Inst. Oswaldo Cruz* **102**, 617–623 (2007). doi: [10.1590/S0074-0276007005000076](https://doi.org/10.1590/S0074-0276007005000076); pmid: 17710307
133. S. B. Halstead, In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody.

- J. Infect. Dis.* **140**, 527–533 (1979). doi: [10.1093/infdis/140.4.527](https://doi.org/10.1093/infdis/140.4.527); pmid: [117061](https://pubmed.ncbi.nlm.nih.gov/117061/)
134. M. Recker *et al.*, Immunological serotype interactions and their effect on the epidemiological pattern of dengue. *Proc. Biol. Sci.* **276**, 2541–2548 (2009). doi: [10.1098/rspb.2009.0331](https://doi.org/10.1098/rspb.2009.0331); pmid: [19369266](https://pubmed.ncbi.nlm.nih.gov/19369266/)
 135. D. S. Burke, A. Nisalak, D. E. Johnson, R. M. Scott, A prospective study of dengue infections in Bangkok. *Am. J. Trop. Med. Hyg.* **38**, 172–180 (1988). pmid: [3341519](https://pubmed.ncbi.nlm.nih.gov/3341519/)
 136. S. B. Halstead, J. S. Porterfield, E. J. O'Rourke, Enhancement of dengue virus infection in monocytes by flavivirus antisera. *Am. J. Trop. Med. Hyg.* **29**, 638–642 (1980). pmid: [6157332](https://pubmed.ncbi.nlm.nih.gov/6157332/)
 137. K. B. Anderson *et al.*, Preexisting Japanese encephalitis virus neutralizing antibodies and increased symptomatic dengue illness in a school-based cohort in Thailand. *PLOS Negl. Trop. Dis.* **5**, e1311 (2011). doi: [10.1371/journal.pntd.0001311](https://doi.org/10.1371/journal.pntd.0001311); pmid: [21991398](https://pubmed.ncbi.nlm.nih.gov/21991398/)
 138. L. M. Paul *et al.*, Dengue virus antibodies enhance Zika virus infection (2016); <http://biorxiv.org/content/early/2016/04/25/050112>.
 139. A. H. Fagbami, S. B. Halstead, N. J. Marchette, K. Larsen, Cross-infection enhancement among African flaviviruses by immune mouse ascitic fluids. *Cytobios* **49**, 49–55 (1987). pmid: [3028713](https://pubmed.ncbi.nlm.nih.gov/3028713/)
 140. L. Wang *et al.*, From mosquitos to humans: Genetic evolution of Zika virus. *Cell Host Microbe* **19**, 561–565 (2016). doi: [10.1016/j.chom.2016.04.006](https://doi.org/10.1016/j.chom.2016.04.006); pmid: [27091703](https://pubmed.ncbi.nlm.nih.gov/27091703/)
 141. C. C. de Melo Freire *et al.*, Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans (2015); <http://biorxiv.org/content/early/2015/11/25/032839>.
 142. A. Fajardo, M. Soñora, P. Moreno, G. Moratorio, J. Cristina, Bayesian coalescent inference reveals high evolutionary rates and diversification of Zika virus populations. *J. Med. Virol.* (2016). doi: [10.1002/jmv.24596](https://doi.org/10.1002/jmv.24596); pmid: [27278855](https://pubmed.ncbi.nlm.nih.gov/27278855/)
 143. A. V. S. Hill, Evolution, revolution and heresy in the genetics of infectious disease susceptibility. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 840–849 (2012). doi: [10.1098/rstb.2011.0275](https://doi.org/10.1098/rstb.2011.0275); pmid: [22312051](https://pubmed.ncbi.nlm.nih.gov/22312051/)
 144. A. L. Brass *et al.*, The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell* **139**, 1243–1254 (2009). doi: [10.1016/j.cell.2009.12.017](https://doi.org/10.1016/j.cell.2009.12.017); pmid: [20064371](https://pubmed.ncbi.nlm.nih.gov/20064371/)
 145. J. S. Mackenzie, D. J. Gubler, L. R. Petersen, Emerging flaviviruses: The spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat. Med.* **10** (Suppl.), S98–S109 (2004). doi: [10.1038/nm1144](https://doi.org/10.1038/nm1144); pmid: [15577938](https://pubmed.ncbi.nlm.nih.gov/15577938/)
 146. J. V. Moreno-Mayer *et al.*, Genome-wide ancestry patterns in Rapanui suggest pre-European admixture with Native Americans. *Curr. Biol.* **24**, 2518–2525 (2014). doi: [10.1016/j.cub.2014.09.057](https://doi.org/10.1016/j.cub.2014.09.057); pmid: [25447991](https://pubmed.ncbi.nlm.nih.gov/25447991/)
 147. W. C. Black 4th *et al.*, Flavivirus susceptibility in *Aedes aegypti*. *Arch. Med. Res.* **33**, 379–388 (2002). doi: [10.1016/S0188-4409\(02\)00373-9](https://doi.org/10.1016/S0188-4409(02)00373-9); pmid: [12234528](https://pubmed.ncbi.nlm.nih.gov/12234528/)
 148. T. Chouin-Carneiro *et al.*, Differential susceptibilities of *Aedes aegypti* and *Aedes albopictus* from the Americas to Zika virus. *PLOS Negl. Trop. Dis.* **10**, e0004543 (2016). doi: [10.1371/journal.pntd.0004543](https://doi.org/10.1371/journal.pntd.0004543); pmid: [26938868](https://pubmed.ncbi.nlm.nih.gov/26938868/)
 149. O. Faye *et al.*, One-step RT-PCR for detection of Zika virus. *J. Clin. Virol.* **43**, 96–101 (2008). doi: [10.1016/j.jcv.2008.05.005](https://doi.org/10.1016/j.jcv.2008.05.005); pmid: [18674965](https://pubmed.ncbi.nlm.nih.gov/18674965/)
 150. M. Aubry *et al.*, Seroprevalence of arboviruses among blood donors in French Polynesia, 2011–2013. *Int. J. Infect. Dis.* **41**, 11–12 (2015). doi: [10.1016/j.ijid.2015.10.005](https://doi.org/10.1016/j.ijid.2015.10.005); pmid: [26482390](https://pubmed.ncbi.nlm.nih.gov/26482390/)
 151. World Health Organization, Current Zika Product Pipeline (WHO, 2016); available at www.who.int/csr/research-and-development/zika-rd-pipeline.pdf.
 152. L. R. Bowman, S. Donegan, P. J. McCall, Is dengue vector control deficient in effectiveness or evidence?: Systematic review and meta-analysis. *PLOS Negl. Trop. Dis.* **10**, e0004551 (2016). doi: [10.1371/journal.pntd.0004551](https://doi.org/10.1371/journal.pntd.0004551); pmid: [26986468](https://pubmed.ncbi.nlm.nih.gov/26986468/)
 153. A. L. Wilson *et al.*, Benefit of insecticide-treated nets, curtains and screening on vector borne diseases, excluding malaria: A systematic review and meta-analysis. *PLOS Negl. Trop. Dis.* **8**, e3228 (2014). doi: [10.1371/journal.pntd.0003228](https://doi.org/10.1371/journal.pntd.0003228); pmid: [25299481](https://pubmed.ncbi.nlm.nih.gov/25299481/)
 154. N. L. Achee *et al.*, A critical assessment of vector control for dengue prevention. *PLOS Negl. Trop. Dis.* **9**, e0003655 (2015). doi: [10.1371/journal.pntd.0003655](https://doi.org/10.1371/journal.pntd.0003655); pmid: [25951103](https://pubmed.ncbi.nlm.nih.gov/25951103/)
 155. R. Patterson, Dr. William Gorgas and his war with the mosquito. *CMAJ* **141**, 596–597, 599 (1989). pmid: [2673502](https://pubmed.ncbi.nlm.nih.gov/2673502/)
 156. F. L. Soper, The elimination of urban yellow fever in the Americas through the eradication of *Aedes aegypti*. *Am. J. Public Health Nations Health* **53**, 7–16 (1963). doi: [10.2105/AJPH.53.1.7](https://doi.org/10.2105/AJPH.53.1.7); pmid: [13978257](https://pubmed.ncbi.nlm.nih.gov/13978257/)
 157. The feasibility of eradicating *Aedes aegypti* in the Americas. *Rev. Panam. Salud Publica* **1**, 68–72 (1997). doi: [10.1590/S1020-49891997000100023](https://doi.org/10.1590/S1020-49891997000100023); pmid: [9128110](https://pubmed.ncbi.nlm.nih.gov/9128110/)
 158. G. Kouri, K. Gustavo, Reemergence of dengue in Cuba: A 1998 epidemic in Santiago de Cuba. *Emerg. Infect. Dis.* **4**, 85–88 (1998). doi: [10.3201/eid0401.980111](https://doi.org/10.3201/eid0401.980111); pmid: [9454562](https://pubmed.ncbi.nlm.nih.gov/9454562/)
 159. E.-E. Ooi, K.-T. Goh, D. J. Gubler, Dengue prevention and 35 years of vector control in Singapore. *Emerg. Infect. Dis.* **12**, 887–893 (2006). doi: [10.3201/eid1206.051210](https://doi.org/10.3201/eid1206.051210); pmid: [16707042](https://pubmed.ncbi.nlm.nih.gov/16707042/)
 160. M. S. Fradin, J. F. Day, Comparative efficacy of insect repellents against mosquito bites. *N. Engl. J. Med.* **347**, 13–18 (2002). doi: [10.1056/NEJMoa011699](https://doi.org/10.1056/NEJMoa011699); pmid: [12097535](https://pubmed.ncbi.nlm.nih.gov/12097535/)
 161. A. Lenhart *et al.*, A cluster-randomized trial of insecticide-treated curtains for dengue vector control in Thailand. *Am. J. Trop. Med. Hyg.* **88**, 254–259 (2013). doi: [10.4269/ajtmh.2012.12.0423](https://doi.org/10.4269/ajtmh.2012.12.0423); pmid: [23166195](https://pubmed.ncbi.nlm.nih.gov/23166195/)
 162. C. H. Calisher *et al.*, Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. *J. Gen. Virol.* **70**, 37–43 (1989). doi: [10.1099/0022-1317-70-1-37](https://doi.org/10.1099/0022-1317-70-1-37); pmid: [2543738](https://pubmed.ncbi.nlm.nih.gov/2543738/)
 163. F. Sievers, D. G. Higgins, Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105–116 (2013). doi: [10.1007/978-1-62703-646-7_6](https://doi.org/10.1007/978-1-62703-646-7_6); pmid: [24170397](https://pubmed.ncbi.nlm.nih.gov/24170397/)
 164. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015). doi: [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300); pmid: [25371430](https://pubmed.ncbi.nlm.nih.gov/25371430/)

ACKNOWLEDGMENTS

We thank M. Kraemer and O. Brady for sharing the maps of the global probability of occurrence of *Aedes* and dengue. We also thank N. Reich, J. Konikoff, and J. Williamson for their help with a preliminary systematic review and analysis that laid the groundwork for this Review.

10.1126/science.aaf8160

RESEARCH ARTICLES

STRUCTURAL BIOLOGY

Structure of the voltage-gated K⁺ channel Eag1 reveals an alternative voltage sensing mechanism

Jonathan R. Whicher and Roderick MacKinnon*

Voltage-gated potassium (K_v) channels are gated by the movement of the transmembrane voltage sensor, which is coupled, through the helical S4-S5 linker, to the potassium pore. We determined the single-particle cryo-electron microscopy structure of mammalian K_v10.1, or Eag1, bound to the channel inhibitor calmodulin, at 3.78 angstrom resolution. Unlike previous K_v structures, the S4-S5 linker of Eag1 is a five-residue loop and the transmembrane segments are not domain swapped, which suggest an alternative mechanism of voltage-dependent gating. Additionally, the structure and position of the S4-S5 linker allow calmodulin to bind to the intracellular domains and to close the potassium pore, independent of voltage-sensor position. The structure reveals an alternative gating mechanism for K_v channels and provides a template to further understand the gating properties of Eag1 and related channels.

Excitable cells, such as neurons and muscle cells, transmit electrical impulses known as action potentials. The action potentials are propagated by voltage-gated ion channels, which function analogously to electrical transistors by gating in response to membrane voltage. Voltage-gated potassium (K_v) channels promote K⁺ efflux upon cell membrane depolarization and thereby return the voltage to the K⁺ Nernst or resting membrane potential. K_v channels are tetramers, and each subunit comprises six transmembrane segments (S1 to S6). The S1 to S4 segments form the voltage-sensor (VS) domain, which contains multiple positively charged residues that move in response to the value of the membrane voltage. The movement of the VS is coupled to the pore (S5-S6) by a linker connecting S4 to S5 (S4-S5 linker). In known K_v structures, as well as voltage-dependent Na⁺ and Ca²⁺ channel structures, the S4-S5 linker is an α helix that runs parallel to the membrane and is positioned above S6 (1–4). In this position, the S4-S5 linker was proposed to act as a lever that performs mechanical work on S6 to open and close the channel when the charged S4 helix is driven through the transmembrane electric field (5).

K_v channels can be divided into 12 subfamilies (K_vs 1 to 12). The Eag family, which includes K_v channels 10 to 12, displays sequence homology to K_v channels 1 to 9, the cyclic nucleotide-gated (CNG) channels, and the hyperpolarization-activated (HCN) channels (6). Members of the Eag family have the S1 to S6 transmembrane

segments, as well as three intracellular domains, an N-terminal Per-ARNT-Sim (PAS) domain, a C-terminal C-linker domain, and a C-terminal cyclic nucleotide-binding homology domain (CNBHD). The different gating properties of the Eag family members can be partially explained by the sequence variations among their PAS domains, which demonstrate its importance in gating (7–11). The C-linker domain and CNBHD are homologous to the intracellular domains of HCN and CNG channels. In HCN and CNG, binding of cyclic nucleotides to the cyclic nucleotide-binding domain (CNBD) is coupled to the pore via the C-linker domain (12). In the Eag family, a portion of the CNBHD sequence occupies the putative cyclic nucleotide-binding site, which prevents binding of and renders the Eag channels insensitive to cyclic nucleotides (13–16). Structures of the PAS domain and CNBHD both separately and in complex provided insights into their role in channel gating (9, 14–16). However, the structures did not provide an understanding of how the domains assemble into tetramers or how these domains interact with transmembrane regions. As a result, there remain many questions as to how the intracellular domains gate the channel and how this gating mechanism compares with that of the other K_v channels and the CNG channels.

K_v10.1, or Eag1, is member of the Eag family that produces a potassium-selective, noninactivating current (17). Activation time of Eag1 is sensitive to hyperpolarization, a property that may be regulated by the interaction between the N terminus of the PAS domain and the S4-S5 linker (7, 8). In addition, Eag1 is inhibited by the binding of calmodulin (CaM) only in the presence of calcium (18, 19). Eag1 is found primarily in the brain, where it localizes to pre-

synaptic terminals and regulates neurotransmitter release (20). Furthermore, Eag1 functions in the differentiation of myoblasts (21) and in the regulation of cell cycle progression (22). Overexpression of Eag1 in mammalian cells induces an oncogenic phenotype, and Eag1 is found overexpressed in tumor tissues from a wide range of cancers (23, 24). Inhibition or knock-down of Eag1 reduces tumor growth (25–29). Therefore, a structure of Eag1 will provide an advance toward an understanding of the gating mechanism of the Eag family, as well as a template for the design of inhibitors that could be used as anticancer therapeutics.

Structure determination of rEag1

We determined the structure of the rat ortholog of Eag1 (rEag1) at 3.78 Å resolution using single-particle cryo-electron microscopy (cryo-EM) (fig. S1). In order to produce stable homogeneous rEag1, it was necessary to truncate the C terminus and purify rEag1 bound to the channel inhibitor CaM (fig. S1). The truncation removed 114 residues (773 to 886) from the unstructured C terminus but retained residues 887 to 962, which are necessary for tetramer formation (30). Truncated rEag1 (rEag1Δ) displayed all the properties of the wild-type (WT) rEag1, including (i) a non-inactivating current with $V_{0.5}$ of 4.9 ± 0.6 mV (SEM) compared with 10 mV for WT Eag1 (31), (ii) activation kinetics that are sensitive to hyperpolarization, and (iii) inhibition by CaM (Fig. 1, A to D).

The overall resolution of the rEag1 cryo-EM density map was 3.78 Å. However, local resolution calculations suggest that the resolution for S1, S2, S4, S5, S6, C-linker, and CNBHD was between 3.3 and 4 Å (fig. S2). In agreement with this calculation, most side chains were observed in these regions of the map, which allowed for de novo model building (fig. S3). The local resolution for PAS, S3, CaM, and regions of the CNBHD (residues 683 to 722) near the CaM binding sites, was between 4 and 6 Å (figs. S2 and S3). As a result, fewer side chains were observed in these regions, but strong main-chain density allowed for accurate model building and placement of crystal structures. The rEag1 model was refined in both real and reciprocal space against one of the independently calculated half maps (working set) (fig. S4). Validation of the model against the other independently calculated half map (free set) and the full map demonstrates good agreement between the model and the maps and the absence of over fitting (fig. S4).

Overall structure of rEag1

Like previously known K_v structures, the transmembrane region of Eag1 has six helical segments (S1 to S6) (1, 2, 32). S1 to S4 form the VS, and S5, the pore helix, and S6 form the potassium pore (Fig. 2). A 40-amino acid turret (residues 378 to 417) between S5 and the pore helix extends ~25 Å out of the membrane and surrounds the extracellular opening of the pore (Fig. 2 and fig. S5). An extended extracellular turret was not observed in previous K_v structures (1, 2, 32), and sequence analysis suggests that, of the eukaryotic K_v channels,

Laboratory of Molecular Neurobiology and Biophysics, The Rockefeller University, Howard Hughes Medical Institute, 1230 York Avenue, New York, NY 10065, USA.

*Corresponding author. Email: mackinn@rockefeller.edu

the Eag family are the only channels that have such an extended extracellular turret (fig. S6). Part of the turret forms an α helix (391 to 399

turret helix), which runs parallel to the membrane and interacts with the pore helix and the loop between the selectivity filter and S6. There

are two glycosylation sites in the turret (N388 and N406) (33), and density was observed for the first *N*-acetyl glucosamine on N388 (figs. S3

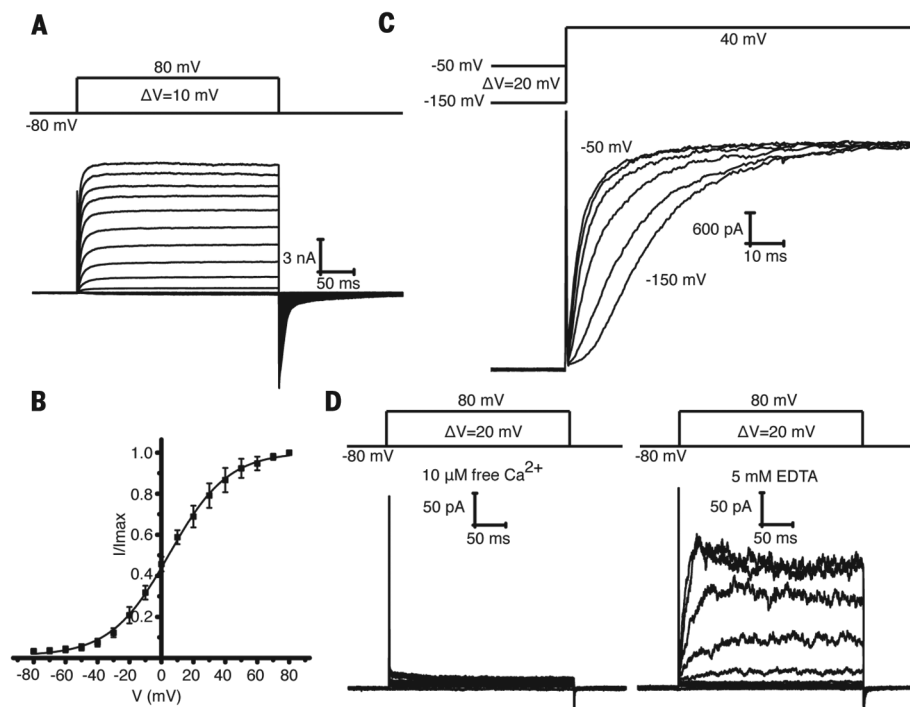


Fig. 1. Functional properties of rEag1Δ. (A) Representative voltage family current trace of rEag1Δ with the voltage-pulse protocol shown above. (B) Tail current (I/I_{\max} , means \pm SD) versus voltage (I - V plot) from (A) was plotted and fit with a Boltzmann function to give a $V_{0.5}$ of 4.9 ± 0.6 mV (SEM) and valence $z = 1.4 \pm 0.05$ (SEM) ($N = 4$). (C) Representative current trace demonstrating that the activation time of rEag1Δ increases after more hyperpolarized (negative) holding potentials. Voltage-pulse protocol is shown above recording. (D) CaM inhibition of rEag1Δ. Inside-out patches from CHO cells expressing rEag1Δ excised in the presence of $10 \mu\text{M}$ free Ca^{2+} (left) and after local perfusion of 5 mM EDTA (right). Voltage-pulse protocol is shown above recording.

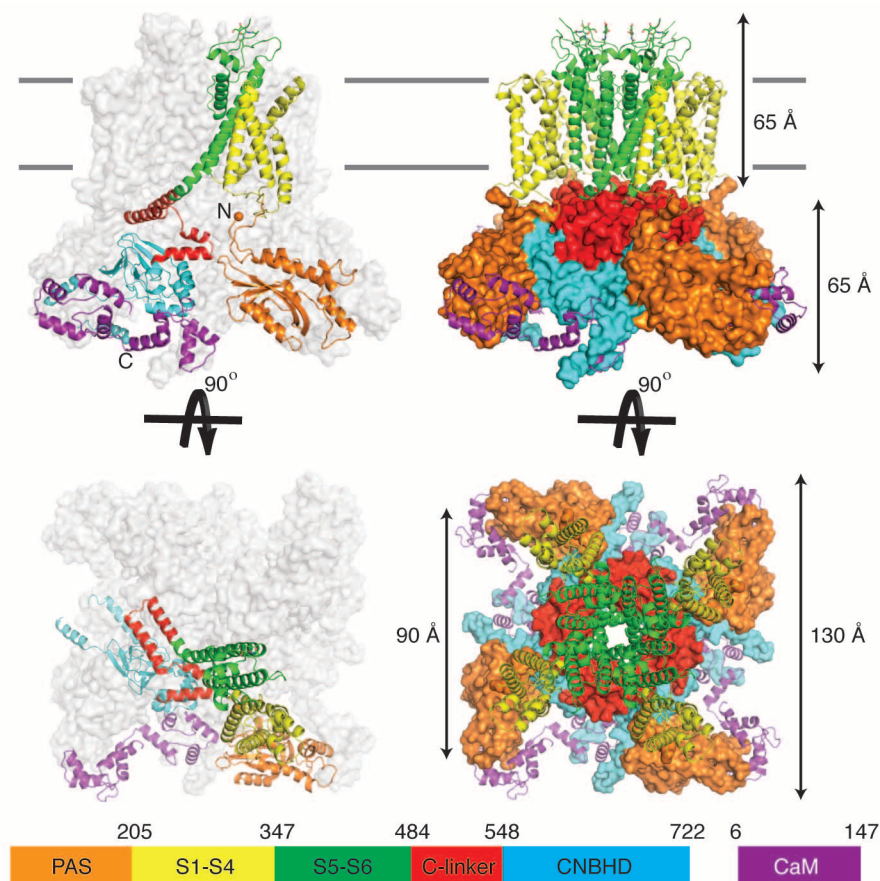


Fig. 2. Model of rEag1Δ bound to CaM. Each domain is colored as indicated, and the cell membrane is shown as gray lines. (Left) One subunit is represented as a cartoon, and the remaining subunits are shown in surface representation. (Right) Trans-membrane domains are represented as a cartoon, and the intracellular domains are represented as a surface. The length and width of trans-membrane and intracellular domains are indicated. The N (orange) and C terminus (cyan) are labeled and shown as spheres. The glycosylation site observed in the turret is shown as sticks with green C, red O, and blue N.

and S5). In this position, the sugar chain may surround the extracellular opening to the pore and prevent binding of inhibitory toxins, which explains why none have been identified for rEag1.

The intracellular domains extend ~65 Å into the cytoplasm. The PAS domains are located at the periphery of the intracellular region, where they interact primarily with the CNBHD from a neighboring subunit and are positioned directly underneath the VS (Fig. 2 and fig. S7). In the context of the full channel, we observed structural elements of the PAS domain not previously seen, including the N-terminal loop (residues 11 to 16), the helix that binds CaM (residues 147 to 157), and a C-terminal loop that connects to S1 of the VS (residues 198 to 213) (figs. S3 and S7). The N terminus of the PAS domain, which influences the rate of voltage-dependent channel opening (activation) and closing (deactivation), is directed toward the VS (Fig. 2 and figs. S3 and S7) (8).

The S6 helix extends into the intracellular region and connects to the C-linker (Fig. 2 and fig. S7). The C-linker has four helices and, as in the HCN channel, the first two helices fold into a helix-turn-helix motif that interacts with the second two helices from the neighboring subunit (12). With this structure, the C-linker forms an intracellular ring directly above the CNBHD ring that can couple the movements of the S6 and CNBHD (Fig. 2 and fig. S7). The CNBHD tetramer displays a quaternary structure similar to that of HCN (12). However, the intracellular vestibule formed by the CNBHD does not contain many negative charges, as is found in many cation channels, which may explain the low conductance of these channels (8 pS) (fig. S7) (34). Like previous CNBHD structures, a portion of the Eag1 sequence occupies the cyclic nucleotide-binding site, which prevents cyclic nucleotide binding (figs. S3 and S7) (14–16).

VS structure

The overall architecture of the Eag1 VS is similar to that of previous K_v structures, but in Eag1, the region of S3 and S4 near the outer membrane leaflet—known as the VS paddle—is shifted toward S2 (fig. S8) (1, 2, 32). The S4 of Eag1 is a long 3_{10} helix that contains six positively charged residues (K327, R330, R333, R336, R339, and K340). In addition, F261 and D264 from S2 and D299 from S3 form a charge-transfer center, which facilitates the movement of the positively charged residues on S4 in response to changes in membrane voltage (35).

Closer inspection of the Eag1 VS reveals three important differences when compared with the $K_v1.2-2.1$ VS (2). The first difference is a 13-residue linker between S2 and S3 (S2-S3 linker) in Eag1 that extends into the cytoplasm (Fig. 3 and figs. S3 and S8). This linker is a conserved feature of the Eag family, and the only other K_v subfamily that contains a similar linker is K_v7 (fig. S6).

The second difference is the register of the positively charged residues on S4 relative to the gating-charge-transfer center (Fig. 3). The S4 of Eag1 and $K_v1.2-2.1$ both have five positively charged residues in equivalent positions (charges

1 to 5; the Gln in position 1 of $K_v1.2-2.1$ is an Arg in $K_v1.2$) (Fig. 3 and fig. S6); however, in the VS of $K_v1.2-2.1$, which is in the open or depolarized conformation, charges 1 to 4 are above (external to) the charge-transfer center, whereas in Eag1 only charges 1 to 3 (K327, R330, and R333) are above the charge-transfer center (2). In Eag1, charge 4 (R336) interacts with D264 and D299 of the charge-transfer center, and although side-chain density was not observed for the charge 5 Arg (R339), it is near D222, which is conserved in the Eag family. We believe this is the depolarized conformation of the Eag1 VS for the following reasons. First, another member of the Eag family, known as hERG or $K_v11.1$, was shown to have a smaller gating charge than $K_v1.2$ (6 to 8 compared with 12 to 14 elementary charge units), which suggests that fewer charges move through the charge-transfer center during hERG gating (36–38). Second, charge-neutralizing mutations indicated that only charges 1 to 3 of hERG are involved in sensing membrane voltage (36). Finally, Eag1 is open at 0 mV, the voltage at which the structure was determined (Fig. 1, A and B). Therefore, in the depolarized conformation of

Eag1, the S4 appears to be displaced one 3_{10} helix turn toward the intracellular membrane when compared with $K_v1.2-2.1$. This allows the S4-S5 linker of Eag1 to interact with the intracellular region of S6 in the up or depolarized conformation (Fig. 4A).

The third difference is the structure of the S4-S5 linker. In Eag1, the S4-S5 linker is a short five-residue loop (Fig. 3A; Fig. 4, A and C; and fig. S6). In $K_v1.2-2.1$, the S4-S5 linker is a 15-residue helix that runs parallel to the membrane (figs. 3B and 4B). This change in the length and structure of the S4-S5 linker has important implications in the architecture of S1 to S6 and channel gating.

Architecture of S1 to S6

The 15-residue helical S4-S5 linker in $K_v1.2-2.1$ structure results in domain-swapped transmembrane segments (VS interacts with the S5 from a neighboring subunit) (Fig. 4B) (1, 2). This domain swap positions the S4-S5 linker of $K_v1.2-2.1$ above S6 from within its own subunit and creates a ring of S4-S5 linker helices that wrap around the S6 helices (1, 2). The five-residue

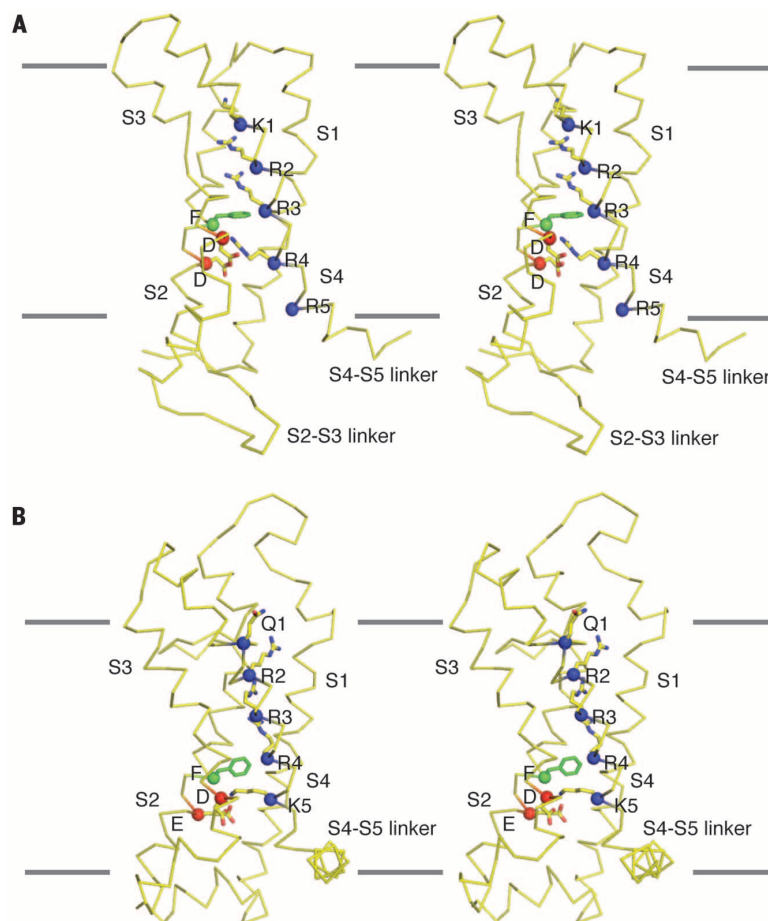


Fig. 3. Structure of VS domains (S1 to S4). Stereoview of the rEag1 (A) and the $K_v1.2-2.1$ (PDB ID 2R9R) (B) VSs (S1 to S4). The five equivalent positions on S4 (blue spheres and sticks with yellow C and blue N; K327, R330, R333, R336, and R339 in rEag1) and the Phe (green spheres and sticks with green C; F261 in rEag1), as well as the negative charges (red spheres and sticks with yellow C and red O; D264 and D299 in rEag1) in the charge transfer. The S2-S3 linker and S4-S5 linker are labeled.

S4-S5 linker of Eag1 cannot span the ~ 20 Å distance between S4 and S5 of the domain-swapped $K_v1.2-2.1$. As a result, the transmembrane segments of Eag1 are not domain-swapped (VS interacts with the S5 from the same subunit) and likely remain in a non-domain swapped conformation domain in all channel states (Fig. 4, A and C). In the non-domain swapped architecture of Eag1, the S4-S5 linker is positioned above the S6 from the neighboring subunit, and the interaction of S1 and S5 makes up the largest interface between the VS and pore (Fig. 4A and fig. S8).

In $K_v1.2-2.1$, the position and structure of the S4-S5 linker, as well as mutational studies on the closely-related Shaker K_v channel, suggest that the S4-S5 linker functions as a mechanical lever to couple movement of the VS to S6 (Fig. 4B) (5, 39). A conserved proline in S6 was proposed to act as a hinge (Pro-Val-Pro hinge), which

allows S6 to move up and down in the membrane in response to movement of the S4-S5 linker to open and close the channel (5, 40). However, the Eag1 S4-S5 linker is a short loop that is not domain swapped and, thus, is not in a position to function as a mechanical lever. Indeed, cutting the S4-S5 linker in Eag1, that is, deleting the covalent link between VS and the pore, produces a functional voltage-gated channel (41). Furthermore, the S6 hinge proline is replaced by G469 in Eag1 (fig. S6), which mutational experiments suggest does not function as a hinge (42). In agreement, S6 of Eag1 is entirely helical with no noticeable kink before the S6 bundle crossing (Fig. 4A). These structural differences suggest that the mechanism of coupling VS movement to pore opening and closing is fundamentally different between Eag1 and $K_v1.2-2.1$. Sequence analysis indicates that all Eag family members

have a short S4-S5 linker and no hinge proline in S6, whereas K_v s 1 to 9 have a longer S4-S5 linker and a hinge proline (Figs. 3 and 4 and fig. S6). Therefore, we predict that all Eag family members will have a non-domain swapped architecture and a gating mechanism similar to Eag1, and K_v s 1-9 will have a domain-swapped architecture and a gating mechanism similar to that of $K_v1.2-2.1$. Why is the mechanism of coupling VS movement to pore opening and closing different between the Eag family and K_v s 1 to 9 and how is movement of the VS coupled to the pore in Eag1? To answer these questions, it is necessary to first understand the mechanism of CaM inhibition and the interaction between the intracellular domains and the VS.

Mechanism of CaM inhibition

CaM binds Eag1 in the presence of Ca^{2+} and inhibits ion conduction (Fig. 1D) (18, 19). In agreement with this, when bound to CaM, the rEag1 pore is closed with a diameter (~ 1 Å between van der Waal surfaces) less than that of hydrated potassium (6 to 8 Å) (Fig. 5A). In this closed conformation, the S6 of rEag1 has an $\sim 55^\circ$ bend directly after the inner helix gate (Q476) in a direction that would tighten the S6 helical bundle (Fig. 5B).

Each Eag1 subunit has three CaM contact regions, which form two binding sites (19). In rEag1, all three regions are occupied by one CaM molecule and thus four CaM molecules are bound to the rEag1 tetramer (Fig. 5, B and C). The two contact regions at the C terminus of the CNBHD form two α helices that act as a claw to grab the CaM C-lobe. The first helix (residues 683 to 695), which has been shown to bind with lower affinity, interacts with the top of the CaM C-lobe near the Ca^{2+} binding sites and the second helix (residues 708 to 722), which has been shown to bind with higher affinity, is bound to the hydrophobic pocket of the CaM C-lobe (Fig. 5, B and C, and fig. S9) (19). The linker helix in CaM that connects the C- and N-lobes breaks and forms a loop from M76 to D80, a segment of high flexibility in CaM (fig. S9) (43). The break in the linker helix allows the hydrophobic core of the CaM N-lobe to bind a helix (residues 147 to 157) on the PAS domain (Fig. 5, B and C, and fig. S9). Note that this PAS domain interacts with the neighboring CNBHD. We propose the modeled orientation of the CaM lobes for two reasons. First, the linker constraints will only allow for this orientation (fig. S9). Second, further classification of the rEag1 particles generated a cryo-EM density map with improved CaM density that shows the position of the linker connecting the CaM lobes (fig. S9).

By binding to the CNBHD and PAS from neighboring subunits, we propose that the CaM acts as a molecular clamp to pull the two domains together and, thereby, alters the orientation of neighboring CNBHD domains (cyan and gray in Fig. 5C). Such a clamping mechanism would translate the CNBHD interacting with the CaM (cyan in Fig. 5C) toward the neighboring PAS domain (Fig. 6A). Because the CNBHD domain

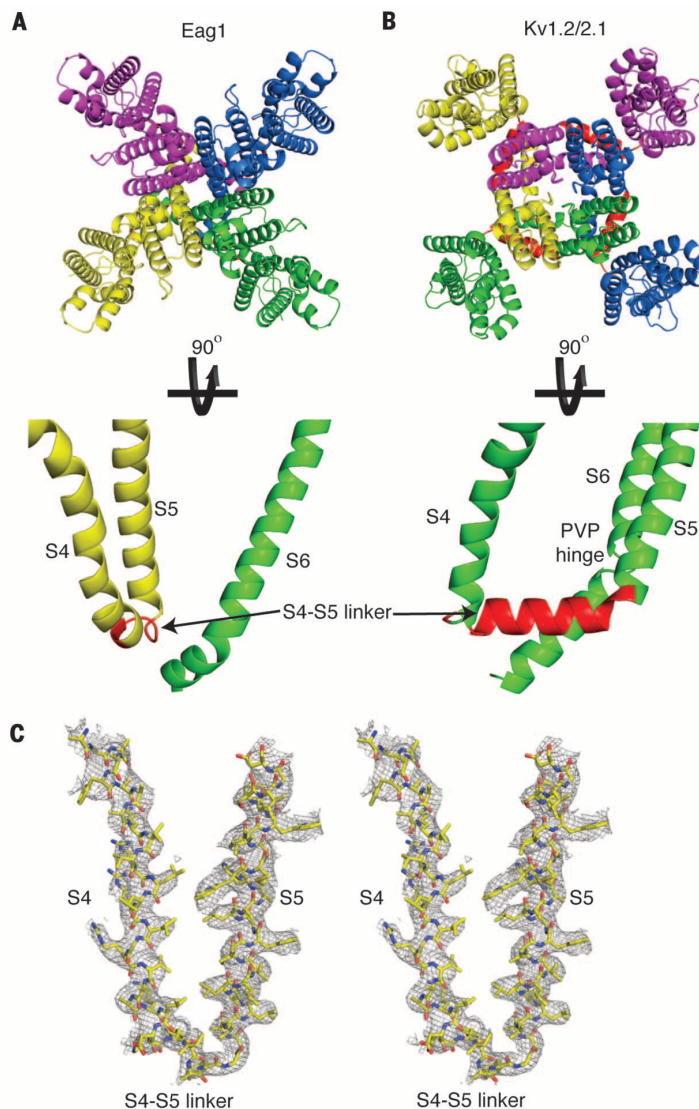


Fig. 4. Architecture of transmembrane domains (S1 to S6). The transmembrane segments (S1 to S6) of rEag (A) and $K_v1.2-2.1$ (B) (PDB ID 2R9R) are shown with each subunit represented as a different color and the S4-S5 linker shown in red. (C) Stereoview of S4, the S4-S5 linker, and S5 segments of rEag1 (shown as sticks with yellow C, red O, and blue N) with cryo-EM density.

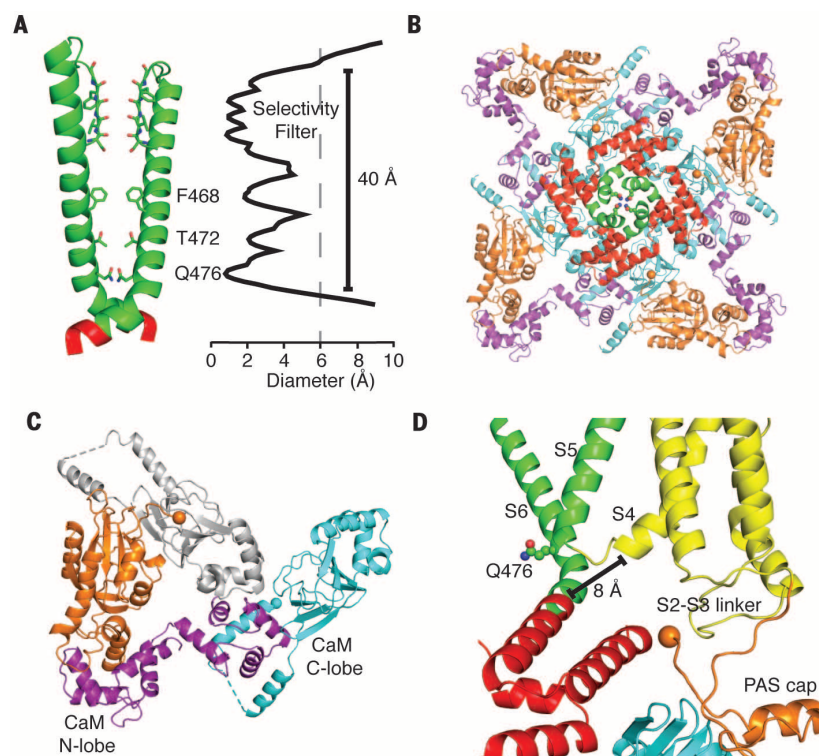


Fig. 5. Intracellular domains and CaM interaction. (A) Two subunits of the tetramer are shown, and sticks (green C, red O, and blue N) indicate the selectivity filter (436 to 441), constriction sites (F468 and T472), and the bundle crossing (Q476) (left). The plot of pore diameter (right) is aligned with the molecular model. At all points, the pore diameter is smaller than that of hydrated potassium (dashed gray line at 6 Å). (B) Organization of rEag1 intracellular domains (extracellular view). Directly C-terminal to the bundle crossing [spheres colored as in (A)], the S6 helix (green) is connected to the C-linker (red). The CaM (purple) is bound to both the CNBHD (cyan) and the PAS (orange). (C) Two helices (683 to 695 and 708 to 722) at the C terminus of the CNBHD (cyan, a dashed line shows the connection between helices) act as a claw to grab the CaM C-lobe (purple). The N-lobe of the same CaM molecule binds to a helix on the PAS domain (147 to 157, orange) that interacts with the neighboring CNBHD (gray). (D) The pore is closed with a 55° bend in S6 in a direction that tightens the helical bundle that forms the intracellular gate [Q476, spheres colored as in (A)]. The VS (yellow) is in an up or depolarized conformation with the S4 directed toward the C-linker at a distance of 8 Å. The N terminus of the PAS domain (orange sphere) and the S2-S3 linker are directed toward the interface between VS and C-linker.

is connected to S6 via the C-linker, we hypothesize that the movement of the CNBHD toward the PAS domain would cause a rotation of both the C-linker and S6 to induce the 55° bend in a direction that tightens the helical bundle that forms the intracellular gate and to close the pore (Fig. 6A). Thus, CaM binding appears to twist the pore helices shut.

Insights into voltage-dependent gating

Functionally, CaM binding induces pore closure independent of voltage (Fig. 1D) (18, 19). The basis for this effect is observed in the structure of Eag1 bound to CaM, as the closed pore is decoupled from the up or depolarized VS (Figs. 3A and 5D). The non-domain swapped architecture of the Eag1, in which the S4-S5 linker is a short loop that forms a weak interaction with S6 from the neighboring subunit (Fig. 4A) rather than a helix that wraps around the S6 helices as in $K_v1.2-2.1$ (Fig. 4B), allows for pore closure by

cytoplasmic domains independent of VS conformation. In this closed conformation, the C terminus of S4 is directed toward the C-linker at a distance of ~8 Å (Fig. 5D). We hypothesize that upon release of CaM, the C-linker and S6 would rotate toward the S4 in a direction that loosens the helical bundle. Such a rotation would relieve the high-energy 55° bend in S6 and open the channel (Fig. 6B).

In the up or depolarized conformation of the VS, the S4-S5 linker interacts with the intracellular portion of S6 and thereby directs the C terminus of S4 toward the C-linker (Fig. 5D). In a down or hyperpolarized conformation of the VS, we hypothesize that the C terminus of S4 will move toward the cytoplasm to interact with the C-linker and will induce a bend in S6 similar to that imposed by CaM binding to close the channel (Fig. 6B). This mechanism, in which S4, rather than the S4-S5 linker, directly interacts with the C-linker to close the channel,

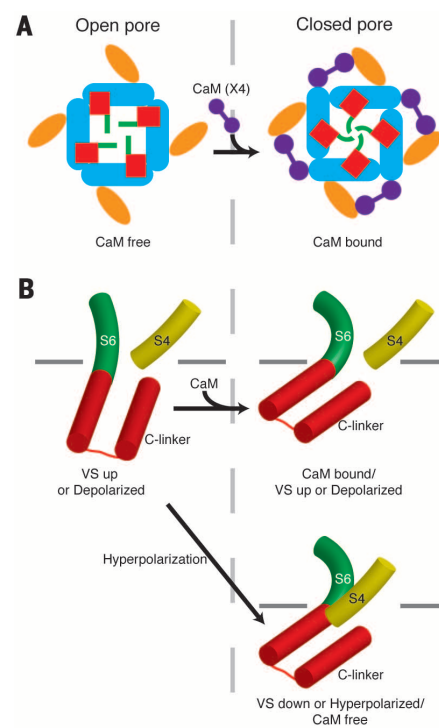


Fig. 6. Proposed mechanism of CaM inhibition and voltage-dependent gating. (A) By binding to the PAS (orange) and CNBHD (cyan), the CaM (purple) molecule seems to pull the CNBHD, which is connected to the S6 (green) via the C-linker (red), toward the PAS domain. This would cause a clockwise rotation (when viewed from the extracellular side of the membrane) of the C-linker and S6 to close the potassium pore. (B) In the structure of rEag1 bound to CaM, the pore is closed with a 55° bend in S6 (green cylinder) in a direction that tightens the helical bundle that forms the intracellular gate, but the VS is in an up or depolarized conformation with the S4 (yellow cylinder) directed toward the C-linker (red cylinders) at a distance of 8 Å (top right, same conformation as Fig. 5D). In an up or depolarized state of the VS in the absence of CaM, the C-linker might rotate in a direction that loosens the helical bundle to relieve the curve in S6 and open the channel (left). In the closed or hyperpolarized state of the VS, the downward movement of the S4 might contact the C-linker to induce a bend in S6 similar to that imposed by CaM binding to close the channel (bottom right). Inner leaflet of the membrane is indicated by gray lines.

would not require a covalent link between S4 and S5 to produce a voltage-dependent channel and, thus, is consistent with previous reports (41). The N terminus of the PAS domain extends from the helical PAS cap (residues 17 to 25) and is directed toward the interface between S4 and C-linker through interactions with the PAS C-terminal loop, the S2-S3 linker, and the CNBHD (Fig. 5D and fig. S7). In this position, the N terminus of the PAS domain may interact with the VS in a down or hyperpolarized conformation to stabilize the closed state and confer

activation time sensitivity to hyperpolarization (8). The S2-S3 linker is also directed toward the S4/C-linker interface and may play a yet-unknown role in channel gating.

Summary

The structure of rEag1 reveals a non-domain swapped architecture of the S1 to S6 that is due to a short five-residue S4-S5 linker. This represents a divergence from the domain-swapped architecture of previous voltage-gated ion channel structures (Fig. 4) (1–4) and suggests a new paradigm for voltage-dependent gating for the Eag family of K_v channels. On the basis of the structure of Eag1, we propose a gating mechanism in which S4 enters the cytoplasm in a down or hyperpolarized state to interact with and induce a rotation of the C-linker and S6 in a direction that tightens the helical bundle to close the channel (Fig. 6B). In the up or depolarized state of the VS, S4 moves into the membrane, which allows the C-linker and S6 to rotate in a direction that loosens the helical bundle and thus relieves the high-energy bend in S6 to open the channel (Fig. 6B).

Two important consequences result from a gating mechanism in which the VS interacts with the cytoplasmic domains to gate the channel. First, this allows the cytoplasmic domains to close the channel independent of VS conformation. This is observed in the structure of Eag1, as binding of CaM to the cytoplasmic domains closes the pore, but the VS is in the up or depolarized conformation. Second, this provides an added level of regulation through the interaction of intracellular domains with the voltage-dependent gating machinery. In Eag1, the N terminus of the PAS domain, which confers sensitivity to hyperpolarization (8), is poised to interact with the S4 and S4-S5 linker in a closed conformation (Fig. 5D).

REFERENCES AND NOTES

1. S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 897–903 (2005).
2. S. B. Long, X. Tao, E. B. Campbell, R. MacKinnon, *Nature* **450**, 376–382 (2007).
3. J. Payandeh, T. Scheuer, N. Zheng, W. A. Catterall, *Nature* **475**, 353–358 (2011).
4. J. Wu et al., *Science* **350**, aad2395 (2015).
5. S. B. Long, E. B. Campbell, R. MacKinnon, *Science* **309**, 903–908 (2005).
6. H. R. Guy, S. R. Durell, J. Warmke, R. Drysdale, B. Ganetzky, *Science* **254**, 730 (1991).
7. M. Ju, D. Wray, *Biochem. Biophys. Res. Commun.* **342**, 1088–1097 (2006).
8. H. Terlau, S. H. Heinemann, W. Stühmer, O. Pongs, J. Ludwig, *J. Physiol.* **502**, 537–543 (1997).
9. J. H. Morais Cabral et al., *Cell* **95**, 649–655 (1998).
10. F. W. Muskett et al., *J. Biol. Chem.* **286**, 6184–6191 (2011).
11. J. Wang, M. C. Trudeau, A. M. Zappia, G. A. Robertson, *J. Gen. Physiol.* **112**, 637–647 (1998).
12. W. N. Zagotta et al., *Nature* **425**, 200–205 (2003).
13. T. I. Brelidze, A. E. Carlson, W. N. Zagotta, *J. Biol. Chem.* **284**, 27989–27997 (2009).
14. T. I. Brelidze, A. E. Carlson, B. Sankaran, W. N. Zagotta, *Nature* **481**, 530–533 (2012).
15. M. J. Marques-Carvalho et al., *J. Mol. Biol.* **423**, 34–46 (2012).
16. Y. Haitin, A. E. Carlson, W. N. Zagotta, *Nature* **501**, 444–448 (2013).

17. J. Ludwig et al., *EMBO J.* **13**, 4451–4458 (1994).
18. R. Schönherr, K. Löber, S. H. Heinemann, *EMBO J.* **19**, 3263–3271 (2000).
19. U. Ziechner et al., *FEBS J.* **273**, 1074–1086 (2006).
20. L. S. Mortensen et al., *J. Physiol.* **593**, 181–196 (2014).
21. P. Bijlenga et al., *J. Physiol.* **512**, 317–323 (1998).
22. L. A. Pardo, A. Brüggemann, J. Camacho, W. Stühmer, *J. Cell Biol.* **143**, 767–775 (1998).
23. L. A. Pardo et al., *EMBO J.* **18**, 5540–5547 (1999).
24. B. Hemmerlein et al., *Mol. Cancer* **5**, 41 (2006).
25. J. R. Agarwal, F. Griesinger, W. Stühmer, L. A. Pardo, *Mol. Cancer* **9**, 18 (2010).
26. F. Mello de Queiroz, G. Suarez-Kurtz, W. Stühmer, L. A. Pardo, *Mol. Cancer* **5**, 42 (2006).
27. B. R. Downie et al., *J. Biol. Chem.* **283**, 36234–36240 (2008).
28. D. Gómez-Varela et al., *Cancer Res.* **67**, 7343–7349 (2007).
29. J. García-Quiroz et al., *PLOS ONE* **7**, e45063 (2012).
30. J. Ludwig, D. Owen, O. Pongs, *EMBO J.* **16**, 6337–6345 (1997).
31. M. Ju, D. Wray, *FEBS Lett.* **524**, 204–210 (2002).
32. Y. Jiang et al., *Nature* **423**, 33–41 (2003).
33. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
34. P. H. Hsu et al., *PLOS ONE* **7**, e41203 (2012).
35. X. Tao, A. Lee, W. Limapichat, D. A. Dougherty, R. MacKinnon, *Science* **328**, 67–73 (2010).
36. M. Zhang, J. Liu, G. N. Tseng, *J. Gen. Physiol.* **124**, 703–718 (2004).
37. S. K. Aggarwal, R. MacKinnon, *Neuron* **16**, 1169–1177 (1996).
38. S. A. Seoh, D. Sigg, D. M. Papazian, F. Bezanilla, *Neuron* **16**, 1159–1167 (1996).

39. Z. Lu, A. M. Klem, Y. Ramu, *Nature* **413**, 809–813 (2001).
40. D. del Camino, M. Holmgren, Y. Liu, G. Yellen, *Nature* **403**, 321–325 (2000).
41. É. Lörinczi et al., *Nat Commun* **6**, 6672 (2015).
42. R. M. Hardman, P. J. Stansfeld, S. Dalibalta, M. J. Sutcliffe, J. S. Mitcheson, *J. Biol. Chem.* **282**, 31972–31981 (2007).
43. G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, A. Bax, *Biochemistry* **31**, 5269–5278 (1992).

ACKNOWLEDGMENTS

We thank M. Ebrahim at the Rockefeller University cryo-EM resource center for help with data collection and J. Chen and members of the MacKinnon laboratory for helpful discussions. This work was supported in part by NIH grant GM43949. J.R.W. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2212-15) and R.M. is an investigator in the Howard Hughes Medical Institute. The low-pass filtered and amplitude-modified three-dimensional cryo-EM density maps for rEag1 have been deposited in the EM Data Bank with accession code EMD-8215. Atomic coordinates for rEag1 have been deposited in the Protein Data Bank under accession code 5K7L. J.R.W. performed functional experiments and expressed, purified, and determined single-particle cryo-EM structure of rEag1. J.R.W. and R.M. designed experiments, analyzed and interpreted results, and wrote the manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/664/suppl/DC1
Materials and Methods
Figs. S1 to S9
References (44–65)

1 April 2016; accepted 22 June 2016
10.1126/science.aaf8070

PHYSICS

Laser spectroscopy of muonic deuterium

Randolf Pohl,^{1,2*} François Nez,³ Luis M. P. Fernandes,⁴ Fernando D. Amaro,⁴ François Biraben,³ João M. R. Cardoso,⁴ Daniel S. Covita,⁵ Andreas Dax,⁶ Satish Dhawan,⁶ Marc Diepold,¹ Adolf Giesen,^{7,8†} Andrea L. Gouvea,⁴ Thomas Graf,⁷ Theodor W. Hänsch,^{1,9} Paul Indelicato,³ Lucile Julien,³ Paul Knowles,^{10‡} Franz Kottmann,¹¹ Eric-Olivier Le Bigot,³ Yi-Wei Liu,¹² José A. M. Lopes,^{4,13} Livia Ludhova,^{10§} Cristina M. B. Monteiro,⁴ Françoise Mulhauser,^{10,1‡} Tobias Nebel,^{14¶} Paul Rabinowitz,¹⁴ Joaquim M. F. dos Santos,⁴ Lukas A. Schaller,¹⁰ Karsten Schuhmann,^{11,8,15} Catherine Schwob,³ David Taqqi,¹⁵ João F. C. A. Veloso,⁵ Aldo Antognini,^{1,11,15} The CREMA Collaboration

The deuteron is the simplest compound nucleus, composed of one proton and one neutron. Deuteron properties such as the root-mean-square charge radius r_d and the polarizability serve as important benchmarks for understanding the nuclear forces and structure. Muonic deuterium μd is the exotic atom formed by a deuteron and a negative muon μ^- . We measured three 2S-2P transitions in μd and obtain $r_d = 2.12562(78)$ fm, which is 2.7 times more accurate but 7.5σ smaller than the CODATA-2010 value $r_d = 2.1424(21)$ fm. The μd value is also 3.5σ smaller than the r_d value from electronic deuterium spectroscopy. The smaller r_d , when combined with the electronic isotope shift, yields a “small” proton radius r_p , similar to the one from muonic hydrogen, amplifying the proton radius puzzle.

Precision spectroscopy of atomic energy levels can be used to determine properties of the nucleus (*I*). Deuterium (*D*), for example, is a heavier isotope of hydrogen (*H*), with a nucleus, the deuteron (*d*), composed of one proton and one neutron (*2*). *D* was dis-

covered through a tiny shift in the Balmer spectral lines of D-enriched hydrogen (*3*). This shift is caused mainly by the mass difference between the proton and the deuteron. Today, the nuclear masses are accurately known from cyclotron frequency measurements in a Penning trap

(*I*), and the measured isotope shift of the 1S-2S transition in H and D (4) determines the (squared) deuteron-proton charge radius difference (5)

$$\delta^{(2)}(\text{H, D}) \equiv r_d^2 - r_p^2 = 3.82007(65) \text{ fm}^2 \quad (1)$$

This is because the wave function of atomic S states is maximal at the origin, where the nucleus resides, and the wave function overlap with the extended nuclear charge distribution reduces the atomic binding energy. Equation 1 links measurements of transition frequencies in H and D. These, together with elastic electron scattering on protons (6) and deuterons (7), determine the Rydberg constant R_∞ , r_p and r_d in the CODATA adjustment of the fundamental physical constants (1).

Muonic atoms are a special class of “exotic” atoms that offer access to nuclear properties with much higher accuracy. In a muonic atom, the nucleus is orbited by one negative muon μ^- , instead of the usual electrons e^- . The muon’s larger mass $m_\mu = 207m_e$ results in a muonic Bohr radius that is smaller than the corresponding electronic Bohr radius by the ratio of reduced masses $m_{\text{red}} = m_\ell m_{\text{nuc}} / (m_\ell + m_{\text{nuc}})$. Here m_ℓ is the mass of the lepton (muon μ^- or electron e^-), and m_{nuc} is the mass of the nucleus. As the Bohr radius shrinks proportionally to $1/m_{\text{red}}$, the overlap of the muon’s wave function with the nuclear charge distribution increases as m_{red}^3 . For μd , $m_{\text{red}} = 196m_e$, and the wave function overlap is $(m_{\text{red}}/m_e)^3 \approx 10^7$ larger in μd than in D. A measurement of the Lamb shift (2P-2S energy difference) in μd is therefore extremely sensitive to the deuteron charge radius r_d .

Our recent measurements of the Lamb shift in muonic hydrogen μp have resulted in a value of the proton charge radius $r_p = 0.84087(39) \text{ fm}$, which is 10 times more accurate, but 4%, or 7σ , smaller (8, 9) than the CODATA-2010 value (1), which is the most recently published CODATA compilation. This so-called “proton radius puzzle”

has questioned the correctness of various experiments or quantum electrodynamics (QED) calculations, the value of the Rydberg constant, our understanding of the proton structure, or the standard model of particle physics (10, 11).

Here we present measurements of the three 2S-2P transitions in μd highlighted in Fig. 1, yielding a precise value of r_d . The principle of the experiment is to form μd atoms in the metastable 2S state (12) and to measure the 2S-2P transitions by pulsed laser spectroscopy. Comparison with theory (13) reveals r_d . The muonic deuterium data presented here were acquired in the same measurement period as the muonic hydrogen data in (8, 9). Independent and reliable calculations of QED (14–17) and nuclear structure effects (18–22) in μd , which are required to interpret the experiment, have recently become available and are summarized in (13).

Measurement of the spectral lines of muonic deuterium

The experiment has been described before (8, 9). In brief, a $5 \times 12 \text{ mm}^2$ beam of low-energy negative muons μ^- (3-keV kinetic energy, average rate 600/s) is stopped in a 20-cm-long target filled with 1 hPa of D_2 gas at 20°C. A pulsed laser system (23, 24) is triggered on the detection of a single arriving muon and provides pulses with an energy of $\sim 0.25 \text{ mJ}$, tunable around a wavelength of $6 \mu\text{m}$, and calibrated against water vapor absorption lines known within a few megahertz (25). A multipass mirror cavity (26) ensures good laser illumination of the muon stop volume. Large-area avalanche photo diodes (27, 28) detect the 2-keV K_α x-rays from the radiative $2\text{P} \rightarrow 1\text{S}$ transition that follows the laser-induced $2\text{S} \rightarrow 2\text{P}$ excitation of μd . The laser frequency is changed every few hours, and the resonances displayed in Fig. 2 are obtained by plotting the number of 2-keV x-rays (normalized to the number of stopped muons) detected in time coincidence with the laser pulse, as a function of laser frequency. On the peak of the resonance, we recorded up to 10 laser-induced x-rays (“events”) per hour with all data reduction cuts (9) applied. The background level of about 2 events per hour originates mainly from misidentified muon decay electrons. About a third of the recorded events are without laser light, providing the expected background level shown as horizontal bands in Fig. 2. The resonances are fitted with a flat background plus a Lorentzian line shape model that takes into account varying laser pulse energies and saturation effects.

The three resonances shown in Fig. 2 are the μd transitions $2\text{S}_{1/2}^{F=3/2} \rightarrow 2\text{P}_{3/2}^{F=5/2}$, $2\text{S}_{1/2}^{F=1/2} \rightarrow 2\text{P}_{3/2}^{F=3/2}$, and $2\text{S}_{1/2}^{F=1/2} \rightarrow 2\text{P}_{1/2}^{F=1/2}$, abbreviated as #1, #2, and #3, respectively. Their positions and uncertainties are

$$\nu_1 = 50816.27 \pm 0.84(\text{stat}) \pm 0.35(\text{syst}) \text{ GHz} \quad (2)$$

$$\nu_2 = 52061.2 \pm 2.0(\text{stat}) \pm 0.35(\text{syst}) \text{ GHz} \quad (3)$$

$$\nu_3 = 52154.1 \pm 2.2(\text{stat}) \pm 0.35(\text{syst}) \text{ GHz} \quad (4)$$

The systematic uncertainties of 0.35 GHz arise from laser frequency fluctuations (8) and Zeeman shifts from a conceivable small admixture of circular polarized light and the 5 T magnetic field of the muon beam line. Line-pulling effects from off-resonant excitation of neighboring levels are negligible (29).

Deuteron charge radius

For the fit of line #1, the Lorentzian width was fixed to the natural radiative line width of $\Gamma = 19.5 \text{ GHz}$ (8, 9), as the freely fitted value $\Gamma = 13.1 \text{ GHz}$ is 2.6σ too small. Both fits agreed on the line center within 0.33 GHz, and the uncertainty quoted in Eq. 2 is the larger one from the fit with fixed width. The difference $\nu_3 - \nu_2 = 92.9 \pm 3.0 \text{ GHz}$ from the fit is in good agreement (1.5σ) with the theoretical value of 88.045 GHz (13). The amplitude of line #3 is larger than zero only with a significance of 4.5σ , but it serves to identify line #2 unambiguously. The alternative—namely, that the left peak in Fig. 2 (bottom) is in fact line #3—is disfavored with 6.9σ significance thanks to the absence of a peak with twice the amplitude $\sim 90 \text{ GHz}$ left of line #2.

Combining the three measured frequencies and using the theoretical 2P fine structure and $2\text{P}_{3/2}$ hyperfine splittings (13), we determine the 2P-2S Lamb shift (LS) and 2S hyperfine splitting (HFS) in μd

$$\Delta E_{\text{LS}}^{\text{exp}} = 202.8785(31)_{\text{stat}}(14)_{\text{syst}} \text{ meV} \quad (5)$$

$$\Delta E_{\text{HFS}}^{\text{exp}} = 6.2747(70)_{\text{stat}}(20)_{\text{syst}} \text{ meV} \quad (6)$$

with total experimental uncertainties of 0.0034 and 0.0073 meV, respectively. The measured 2S

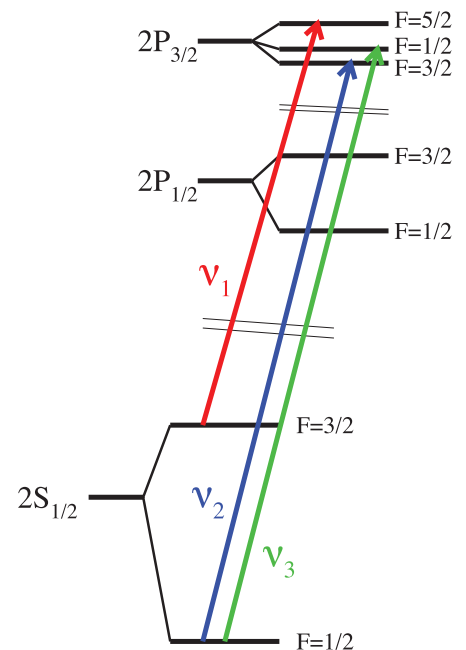


Fig. 1. $n = 2$ levels in muonic deuterium. The order of the $2\text{P}_{3/2}$ sublevels is changed by the nuclear quadrupole moment (13). The three measured transitions are indicated.

¹Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany. ²Johannes Gutenberg-Universität Mainz, QUANTUM, Institut für Physik & Exzellenzcluster PRISMA, 55099 Mainz, Germany. ³Laboratoire Kastler Brossel, UPMC-Sorbonne Universités, CNRS, Ecole Normale Supérieure-PSL Research University, Collège de France, 75005 Paris, France. ⁴LIPPhys, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal. ⁵IN, Departamento de Física, Universidade de Aveiro, 3810-193 Aveiro, Portugal. ⁶Physics Department, Yale University, New Haven, CT 06520-8121, USA. ⁷Institut für Strahlwerkzeuge, Universität Stuttgart, 70569 Stuttgart, Germany. ⁸Dausinger + Giesen GmbH, Rotbühlstrasse 87, 70178 Stuttgart, Germany. ⁹Ludwig-Maximilians-Universität, Munich, Germany. ¹⁰Département de Physique, Université de Fribourg, 1700 Fribourg, Switzerland. ¹¹Institute for Particle Physics, ETH Zurich, 8093 Zurich, Switzerland. ¹²Physics Department, National Tsing Hua University, Hsinchu 300, Taiwan. ¹³Instituto Politécnico de Coimbra, ISEC, 3030-199, Portugal. ¹⁴Department of Chemistry, Princeton University, Princeton, NJ 08544-1009, USA. ¹⁵Paul Scherrer Institute, 5232 Villigen-PSI, Switzerland.

*Corresponding author. Email: pohl@uni-mainz.de †Present address: Deutsches Zentrum für Luft- und Raumfahrt e.V. in der Helmholtz-Gemeinschaft, 70569 Stuttgart, Germany. ‡Present address: LogrusData, Vienna, Austria. §Present address: Forschungszentrum Jülich IKP-2 and RWTH Aachen University, Germany. ||Present address: International Atomic Energy Agency, Vienna, Austria. ¶Present address: Honeywell Process Solutions Inc, 500 Brooksbank Avenue, North Vancouver BC V7J 3S4, Canada.

HFS is in excellent agreement with the theoretical value, $\Delta E_{\text{HFS}}^{\text{theo}} = 6.2791(50)$ meV (13).

The Lamb shift in μd is extraordinarily sensitive (13) to the root mean square (RMS) deuteron charge radius

$$\Delta E_{\text{LS}}^{\text{theo}} = 228.7766(10) \text{ meV} + \Delta E_{\text{LS}}^{\text{TPE}} \quad (7)$$

$$-6.1103(3) r_d^2 \text{ meV/fm}^2$$

where

$$\Delta E_{\text{LS}}^{\text{TPE}}(\text{theo}) = 1.7096(200) \text{ meV} \quad (8)$$

is the deuteron polarizability contribution (13) from two-photon exchange (TPE), recently calculated

with good accuracy (18–22). The charge radius effect in Eq. 7 contributes as much as 14% to the 2P–2S Lamb shift, which explains the excellent sensitivity of our measurement to r_d . We obtain r_d from equating Eqs. 5 and 7, and using Eq. 8, which yields

$$r_d(\mu\text{d}) = 2.12562(13)_{\text{exp}}(77)_{\text{theo}} \text{ fm} \quad (9)$$

where the theory uncertainty is almost exclusively from $\Delta E_{\text{LS}}^{\text{TPE}}$ (Eq. 8). This radius is in 7.5σ disagreement with the CODATA value (1), which is the best estimate of the deuteron radius obtained from precision spectroscopy of H

and D and electron scattering on protons and deuterons,

$$r_d(\text{CODATA}) = 2.1424(21) \text{ fm} \quad (10)$$

(see Fig. 3). We are hence faced with the fact that precision determinations of the Lamb shift in both μp and μd , from a total of five measured resonances, each show a $\geq 7\sigma$ discrepancy to the predictions based on fundamental physical constants from the self-consistent CODATA world average (1), carefully checked QED calculations (13, 30), and physics within the standard model.

The CODATA deuteron radius r_d is tightly linked to the CODATA proton radius r_p , by virtue of Eq. 1. However, as detailed in (31), we have deduced a deuteron charge radius considering spectroscopy data in regular deuterium alone—i.e., without relying on the value of the proton radius. This yields a value of

$$r_d(\text{D spectroscopy}) = 2.1415(45) \text{ fm} \quad (11)$$

in excellent agreement with the CODATA value, but 3.5σ larger than the value obtained here from muonic deuterium (see Fig. 3, blue point, “D spectroscopy”).

This distinct 3.5σ discrepancy between the atomic physics determinations of r_d from D and μd is almost as severe as the 4.0σ atomic physics discrepancy between the r_p values from H spectroscopy [see (1), table XXXVIII, adjustment 8] and μp (see Fig. 4). These two discrepancies are independent, as explained in (31).

The difference between the deuteron radii from the spectroscopy of electronic and muonic deuterium is only 0.017 fm, or 0.8%. Thus, even though the deuteron charge radius $r_d(\text{e-d scatt.}) = 2.130(10)$ fm, extracted from elastic electron-deuteron scattering (7), is accurate to 0.5%, it is unfortunately not accurate enough to distinguish between the values from μd and CODATA.

Proton and deuteron radius puzzle

Many attempts to explain the proton radius discrepancy exist (10, 11). Our muonic deuterium result provides fresh insight, as the so-called “proton radius puzzle” is in fact not limited to the proton; there is a distinct deuteron radius puzzle. Using $r_d(\text{CODATA})$ in Eq. 7 yields a Lamb shift that is $\epsilon_{\text{LS}}(\mu\text{d}) = 0.438(59)$ meV smaller than the measured value, Eq. 5, and hence resonance frequencies that are ~ 104 GHz smaller than observed (Fig. 2). The $\epsilon_{\text{LS}}(\mu\text{d})$ is even somewhat larger than the proton radius discrepancy $\epsilon_{\text{LS}}(\mu\text{p}) = 0.329(47)$ meV between the LS we observed in μp and the one calculated with the CODATA value of r_p (9).

The ratio of discrepancies in μd and μp , $\epsilon_{\text{LS}}(\mu\text{d})/\epsilon_{\text{LS}}(\mu\text{p}) = 1.3(2)$ is in agreement with the ratio of the wave-function overlap from the reduced mass ratio, $[m_{\text{red}}(\mu\text{d})/m_{\text{red}}(\mu\text{p})]^3 = 1.17$. Such a scaling is expected for several beyond-standard model (BSM) physics scenarios (10, 11, 32–34), where a new force between muons and protons is

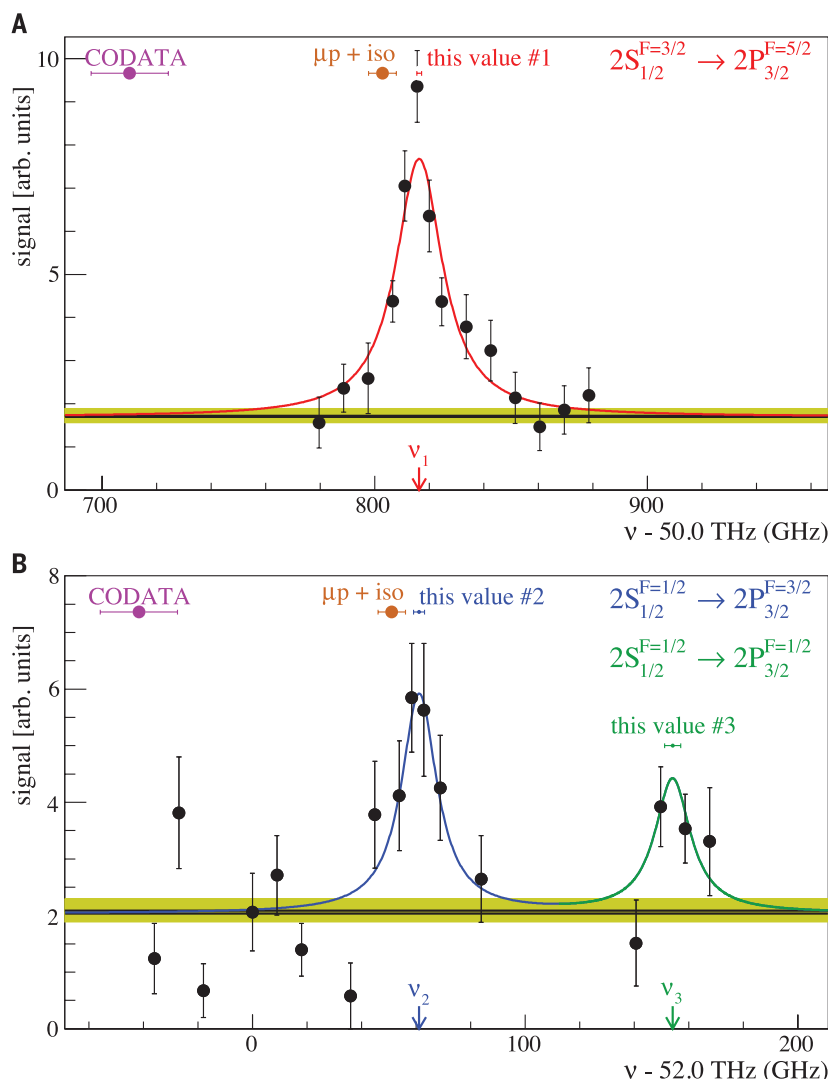
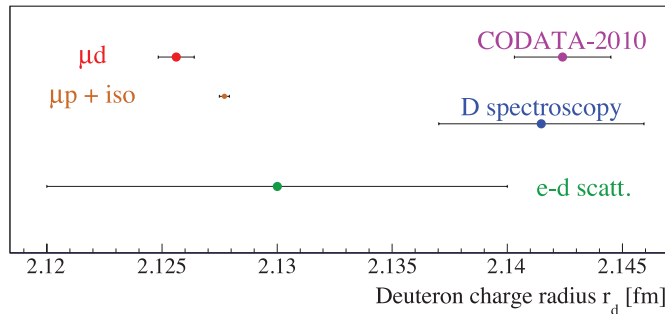


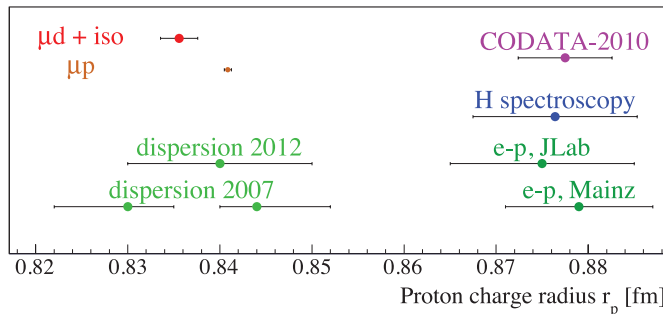
Fig. 2. Three measured resonances in muonic deuterium. The resonances are labeled #1 (A), and #2 and #3 (B). The signal (y axis) is “normalized number of events” as described in (8). Predicted resonance positions are shown based on Eqs. 7 and 8: The CODATA-2010 deuteron radius (pink, Eq. 10) would correspond to ~ 104 GHz lower resonance positions, which is a difference of 7.5σ . The “expected” deuteron radius Eq. 13, (“ μp + iso,” brown) obtained by combining the proton radius from muonic hydrogen (9) and the electronic isotope shift (“iso”), Eq. 1, is consistent with the observed resonance positions within $\sim 2.6\sigma$. The top and bottom panel’s data were recorded in 1 week and 2 days, respectively. As an example, the three highest points around the peak of resonance #1 contain a total of 260 events, recorded in 21 hours.

Fig. 3. Deuteron radii.

Our value Eq. 9 (“ μd ,” red) has a 7.5σ discrepancy with the CODATA-2010 value (1), but is within 2.6σ of the smaller “expected” value Eq. 13 (“ $\mu p + iso$,” brown, see text), obtained by combining the proton radius from muonic hydrogen (9) and the electronic isotope shift, Eq. 1. The value from laser spectroscopy of electronic deuterium (“D spectroscopy,” blue, Eq. 11) is obtained as detailed in (31) and is 3.5σ larger than the μd value. The world average from elastic e-d scattering (7) (“e-d scatt.,” green) is also shown.

**Fig. 4. Proton radii.**

Shown are the values deduced from muonic hydrogen (9) (“ μp ,” brown) and from muonic deuterium (Eq. 9, “ $\mu d + iso$,” red), which both differ by $>7\sigma$ from the CODATA-2010 value (1). Many other determinations exist, and we highlight the values from spectroscopy of H (but not D), from CODATA-2010 [(1), table XXXVIII, adjustment 8] (“H spectroscopy,” blue); elastic electron-proton scattering (dark green) from (6) and (47); and electron scattering data analyses based on dispersion relations (light green), both less recent (48) and more recent (49) than the μp value (8, 9). Many more values exist (50).



responsible for the change in the observed LS and can at the same time explain the long-standing $\sim 3.5\sigma$ discrepancy in the muon g-2 value (35). In these models, the coupling of such a new force to neutrons must be negligible to fulfill other experimental constraints. The same scaling is also expected for explanations based on an unexpectedly large TPE contribution to proton polarizability (36), or an effect of a “sea of leptons” inside the proton (37, 38).

Before resorting to BSM solutions, however, one should investigate what it would take to “solve” the two discrepancies within SM physics. As noted before (8–11), and explained in more detail in (31), the reconciliation of electronic and muonic spectroscopy data still requires rather drastic measures.

For one, the CODATA Rydberg constant could be wrong by $\sim 7\sigma$ —for example, because of a yet-undiscovered, common systematic effect in the most precise measurements of transitions from the 2S to the 8S, 8D, and 12D states in H and D (39). Such a change of R_∞ would shift the proton radius from H to the smaller μp value (8–10, 31). It would also bring the deuteron radius from D to within $\sim 2.5\sigma$ of the μd value (see below).

Alternatively, the QED theory of the Lamb shift in electronic H and D could be missing a large

contribution of ~ 110 kHz, which corresponds to about 44 times the claimed theory uncertainty (44σ) of ~ 2.5 kHz. Such a missing QED contribution would bring the charge radii from H and D spectroscopy into agreement with their muonic counterparts, without changing the Rydberg constant (31).

Third, a systematic shift of all spectroscopic muonic measurements by 140σ (corresponding to 80 GHz in μp and 104 GHz in μd), or a missing theory term in the Lamb shift of muonic atoms that accounts for the missing ϵ_{LS} , could be the source of the discrepancies. This theory error would correspond to 160σ in μp , and 22σ in μd , where the uncertainty of the TPE contribution is about 10 times larger (13). The claimed uncertainty of the pure QED (i.e., non-TPE) contributions in μd (μp) is about 440 (220) times smaller than the ϵ_{LS} (13).

Neither a shift of R_∞ by 7σ from the CODATA value, nor a change of the LS in H and D by ~ 110 kHz, will, however, appreciably affect Eq. 1 (5). Hence, we can proceed and draw conclusions from the fact that the muonic isotope shift

$$\delta^{(2)}(\mu p, \mu d) = 3.81120(339) \text{ fm}^2 \quad (12)$$

is compatible within 2.6σ with the “electronic” isotope shift, Eq. 1, but five times less accurate.

The absolute values of r_p and r_d from the muonic 2S-2P measurements are thus roughly consistent with the size difference from the electronic 1S-2S measurement (4, 5), Eq. 1.

The dominant source of uncertainty in Eq. 12 is the calculated TPE contribution (Eq. 8), whose effect on the uncertainty of r_d from μd , Eq. 9, is about six times larger than the experimental uncertainty. Hence, we are tempted to ascribe the remaining 2.6σ discrepancy between the electronic and muonic isotope shift to the TPE contribution to the LS in μd .

We can thus use the muonic proton radius from μp (9), $r_p(\mu p) = 0.84087(39)$ fm, and the electronic isotope shift, Eq. 1, to obtain a precise value of the deuteron charge radius in an indirect way. The resulting value

$$r_d(\mu p + iso) = 2.12771(22) \text{ fm} \quad (13)$$

was given in (9) and is indicated as “ $\mu p + iso$ ” in Figs. 2 and 3. It is the most accurate value of the deuteron RMS charge radius and is independent of the TPE contribution in μd .

Using this “expected” deuteron radius from Eq. 13 in the theory expression for the LS in μd , Eq. 7, yields an experimental value for the TPE contribution to the LS in μd

$$\Delta E_{LS}^{TPE}(\text{exp}) = 1.7638(68) \text{ meV} \quad (14)$$

from the measured LS in Eq. 5. It is 2.6σ larger than the calculated value, Eq. 8, but three times more accurate, making it a benchmark for ab initio calculations of the deuteron (2, 19, 20, 22) or analysis of virtual Compton scattering data (21).

In a similar manner, we determine the experimental value of the polarizability—i.e., the inelastic part of the TPE contribution to the 2S-HFS—using our measured HFS, Eq. 6, Eq. 42 of (13), and the Zemach radius of the deuteron $r_Z = 2.593(16)$ fm from (40). We obtain

$$\Delta E_{HFS}^{\text{pol}}(\text{exp}) = 0.2178(74) \text{ meV} \quad (15)$$

where the experimental uncertainty is by far the dominant one. This agrees with the theoretical value $\Delta E_{HFS}^{\text{pol}}(\text{theo}) = 0.2226(49)$ meV, which has been calculated only recently (16).

Finally, we note that the reasoning that leads to Eq. 13 can of course be inverted. Using the measured muonic deuteron charge radius, Eq. 9, and the electronic isotope shift, Eq. 1, we obtain a new value for the proton radius

$$r_p(\mu d + iso) = 0.8356(20) \text{ fm} \quad (16)$$

confirming the “small” proton charge radius from muonic hydrogen (8, 9), further amplifying the “proton radius puzzle” (10, 11) (see Fig. 4).

Ultimately, only new experiments can shed more light on the proton and deuteron radius discrepancies. A lot of activity exists in elastic electron scattering (41, 42), with the hope for refined values of r_p and r_d . Muon scattering on the proton will be able to check the BSM hypothesis (43). Moreover,

several atomic physics measurements are underway to verify and improve the Rydberg constant and the proton and deuteron radius from regular (electronic) hydrogen and deuterium (44–46).

REFERENCES AND NOTES

1. P. J. Mohr, B. N. Taylor, D. B. Newell, *Rev. Mod. Phys.* **84**, 1527–1605 (2012).
2. R. Machleidt, *Phys. Rev. C Nucl. Phys.* **63**, 024001 (2001).
3. H. C. Urey, F. G. Brickwedde, G. M. Murphy, *Phys. Rev.* **39**, 164–165 (1932).
4. C. G. Parthey *et al.*, *Phys. Rev. Lett.* **104**, 233001 (2010).
5. U. D. Jentschura *et al.*, *Phys. Rev. A* **83**, 042505 (2011).
6. J. C. Bernauer *et al.*, *Phys. Rev. Lett.* **105**, 242001 (2010).
7. I. Sick, D. Trautmann, *Nucl. Phys. A* **637**, 559–575 (1998).
8. R. Pohl *et al.*, *Nature* **466**, 213–216 (2010).
9. A. Antognini *et al.*, *Science* **339**, 417–420 (2013).
10. R. Pohl, R. Gilman, G. A. Miller, K. Pachucki, *Annu. Rev. Nucl. Part. Sci.* **63**, 175–204 (2013).
11. C. E. Carlson, *Prog. Part. Nucl. Phys.* **82**, 59–77 (2015).
12. M. Diepold *et al.*, *Phys. Rev. A* **88**, 042520 (2013).
13. J. J. Krauth *et al.*, *Ann. Phys. (N.Y.)* **366**, 168–196 (2016).
14. A. A. Krutov, A. P. Martynenko, *Phys. Rev. A* **84**, 052514 (2011).
15. E. Borie, *Ann. Phys. (N.Y.)* **327**, 733–763 (2012).
16. R. N. Faustov, A. P. Martynenko, G. A. Martynenko, V. V. Sorokin, *Phys. Rev. A* **90**, 012520 (2014).
17. R. N. Faustov, A. P. Martynenko, G. A. Martynenko, V. V. Sorokin, *Phys. Rev. A* **92**, 052512 (2015).
18. K. Pachucki, *Phys. Rev. Lett.* **106**, 193007 (2011).
19. J. L. Friar, *Phys. Rev. C Nucl. Phys.* **88**, 034003 (2013).
20. O. Hernandez, C. Ji, S. Bacca, N. Nevo Dinur, N. Barnea, *Phys. Lett. B* **736**, 344–349 (2014).
21. C. E. Carlson, M. Gorchtein, M. Vanderhaeghen, *Phys. Rev. A* **89**, 022504 (2014).
22. K. Pachucki, A. Wienczek, *Phys. Rev. A* **91**, 040503 (2015).
23. A. Antognini *et al.*, *IEEE J. Quantum Electron.* **45**, 993–1005 (2009).
24. A. Antognini *et al.*, *Opt. Commun.* **253**, 362–374 (2005).
25. L. S. Rothman *et al.*, *J. Quant. Spectrosc. Radiat. Transf.* **110**, 533–572 (2009).
26. J. Vogelsang *et al.*, *Opt. Express* **22**, 13050–13062 (2014).
27. L. Ludhova *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **540**, 169–179 (2005).
28. L. M. P. Fernandes *et al.*, *J. Instrum.* **2**, P08005 (2007).
29. P. Amaro *et al.*, *Phys. Rev. A* **92**, 022514 (2015).
30. A. Antognini *et al.*, *Ann. Phys. (N.Y.)* **331**, 127–145 (2013).
31. R. Pohl *et al.*, <https://arxiv.org/abs/1607.03165> (2016).
32. D. Tucker-Smith, I. Yavin, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **83**, 101702 (2011).
33. B. Batell, D. McKeen, M. Pospelov, *Phys. Rev. Lett.* **107**, 011803 (2011).
34. C. E. Carlson, B. C. Rislow, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **86**, 035013 (2012).
35. F. Jegerlehner, A. Nyffeler, *Phys. Rep.* **477**, 1–110 (2009).
36. G. A. Miller, *Phys. Lett. B* **718**, 1078–1082 (2013).
37. U. D. Jentschura, *Phys. Rev. A* **88**, 062514 (2013).
38. G. A. Miller, *Phys. Rev. C Nucl. Phys.* **91**, 055204 (2015).
39. F. Biraben, *Eur. Phys. J. Spec. Top.* **172**, 109–119 (2009).
40. J. Friar, I. Sick, *Phys. Lett. B* **579**, 285–289 (2004).
41. A. Gasparian, *EPJ Web Confer.* **73**, 07006 (2014).
42. M. Mihovilović, H. Merkel, A. Weber, *EPJ Web Confer.* **81**, 01009 (2014).
43. R. Gilman, *AIP Conf. Proc.* **1563**, 167–170 (2013).
44. A. Beyer *et al.*, *Ann. Phys. (Berlin)* **525**, 671–679 (2013).
45. A. C. Vutha *et al.*, *Bull. Am. Phys. Soc.* **57**, Q1138 (2012).
46. S. Galtier *et al.*, *J. Phys. Chem. Ref. Data* **44**, 031201 (2015).
47. X. Zhan *et al.*, *Phys. Lett. B* **705**, 59–64 (2011).
48. M. A. Belushkin, H.-W. Hammer, U.-G. Meissner, *Phys. Rev. C Nucl. Phys.* **75**, 035202 (2007).
49. I. T. Lorenz, U.-G. Meissner, H.-W. Hammer, Y. B. Dong, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **91**, 014023 (2015).
50. A. Antognini, <https://arxiv.org/abs/1512.01765> (2015).

ACKNOWLEDGMENTS

We thank E. Borie for the calculations that made this measurement possible; I. Sick for insightful discussions; L.M. Simons and B. Leoni for setting up the cyclotron trap; R. Rosenfelder and C. Hoffman for support; H. Brückner, K. Linner, W. Simon, O. Huot and Z. Hochman for technical support; P. Maier-Komor, K. Nacke, M. Horisberger, A. Weber, L. Meier, and J. Hehner for thin foils and

windows; N. Schlumpf, U. Hartmann, and M. Gaspar for electronics; S. Spielmann-Jaeggli and L. Carroll for optical measurements; Ch. Parthey and M. Herrmann for their help; Th. Udem for insightful discussions; the MEG-collaboration for a share of valuable beam-time; and A. Voss, B. Weichelt and J. Fruechtenicht for the loan of a laser pump diode. We acknowledge the essential contributions of H. Hofer and V. W. Hughes in the initial stages of the experiment and thank K. Kirch for his continuous support. We also thank the PSI accelerator division, the Hallendienst, the workshops at PSI, MPQ, and Fribourg, and other support groups for their valuable help. We acknowledge support from the European Research Council (ERC StG. 279765), the Max Planck Society and the Max Planck Foundation, the Swiss National Science Foundation (project 200020-100632, 200021L_138175, 200020_159755, 200021_165854) and the Swiss Academy of Engineering Sciences, the BQR de l'UFR de physique fondamentale et appliquée de

l'Université Pierre et Marie Curie- Paris 6, the program PAI Germaine de Staël no. 07819NH du ministère des affaires étrangères France, the Fundação para a Ciência e a Tecnologia (Portugal) and FEDER (project PTDC/FIS/102110/2008 and grants SFRH/BPD/46611/2008, SFRH/BPD/74775/2010, and SFRH/BPD/76842/2011), Deutsche Forschungsgemeinschaft (DFG) GR 3172/9-1 within the D-A-CH framework, and Ministry of Science and Technology, Taiwan, no. 100-2112-M-007-006-MY3. P.I. acknowledges support by the “ExtreMe Matter Institute, Helmholtz Alliance HA216/EMMI.” Reasonable requests for sharing the data should be addressed to R.P. All authors contributed substantially to this work.

13 January 2016; accepted 20 July 2016
10.1126/science.aaf2468

EXTRASOLAR PLANETS

Direct imaging discovery of a Jovian exoplanet within a triple-star system

Kevin Wagner,^{1*} Dániel Apai,^{1,2} Markus Kasper,³ Kaitlin Kratter,¹ Melissa McClure,³ Massimo Robberto,^{4,5} Jean-Luc Beuzit^{6,7}

Direct imaging allows for the detection and characterization of exoplanets via their thermal emission. We report the discovery via imaging of a young Jovian planet in a triple-star system and characterize its atmospheric properties through near-infrared spectroscopy. The semimajor axis of the planet is closer relative to that of its hierarchical triple-star system than for any known exoplanet within a stellar binary or triple, making HD 131399 dynamically unlike any other known system. The location of HD 131399Ab on a wide orbit in a triple system demonstrates that massive planets may be found on long and possibly unstable orbits in multistar systems. HD 131399Ab is one of the lowest mass (4 ± 1 Jupiter masses) and coldest (850 ± 50 kelvin) exoplanets to have been directly imaged.

Thousands of planets around other stars have been discovered (1, 2), revealing a greater diversity than predicted by traditional planet formation models based on the solar system.

Extreme examples are planets within binary and multiple-star systems, which form and evolve in variable radiation and gravitational fields. Direct imaging allows for the detection and spectroscopic characterization of long-period giant planets, thus enabling constraints to be placed on planet formation models via predictions of planet population statistics and atmospheric properties (3). However, most direct imaging surveys have traditionally excluded visual binary or multiple systems whose separations are less than a few hundred astronomical units (AUs). These exclusions are based on the assumption that such planetary systems would either be disrupted or never

form, as well as the increased technical complexity of detecting a planet among the scattered light of multiple stars. As a result of this observational bias, most directly imaged exoplanets have been found around single stars.

Because multistar systems are as numerous as single stars (4), building a complete census of long-period giant planets requires investigation of both configurations. In principal, planets on wide orbits (detectable by direct imaging) might arise more frequently in multistar systems because of planet-planet or planet-star interactions (5, 6). Such interactions could even produce planets on chaotic orbits that wander between the stars (7, 8). To investigate the frequency of long-period giant planets both around single stars and in multistar systems, we are using the Very Large Telescope (VLT) and the Spectro-Polarimetric High-Contrast Exoplanet Research instrument [SPHERE (9)] to sample a population of ~100 young single and multiple A-type stars in the nearby Upper Scorpius-Centaurus-Lupus association. Here we report the discovery of the first planet detected in our ongoing survey and the widest-orbit planet within a multistar system.

Observations and discovery of HD 131399Ab

HD 131399 (also known as HIP72940) is a triple system (10) in the 16 ± 1 -million-year-old Upper

¹Department of Astronomy and Steward Observatory, The University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA. ²Lunar and Planetary Laboratory, The University of Arizona, 1640 East University Boulevard, Tucson, AZ 85718, USA. ³European Southern Observatory (ESO), Karl Schwarzschild-Strasse 2, D-85748 Garching, Germany.

⁴Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA. ⁵Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA. ⁶Université Grenoble Alpes, Institut de Planétologie et d'Astrophysique de Grenoble (IPAG), F-38000 Grenoble, France. ⁷CNRS, IPAG, F-38000 Grenoble, France.

*Corresponding author. Email: kwagner@as.arizona.edu

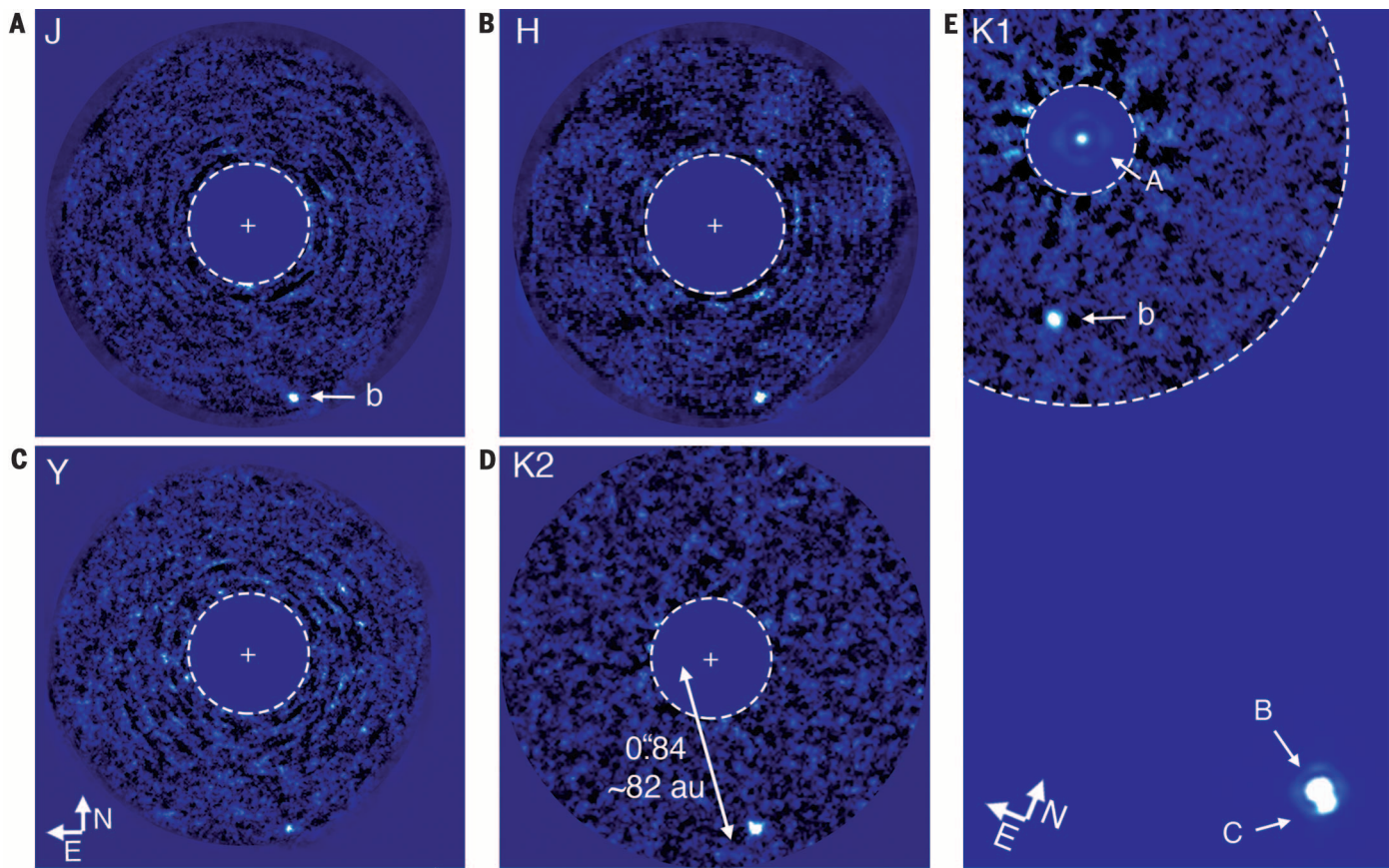


Fig. 1. Near-infrared VLT-SPHERE images of HD 131399Ab and the hierarchical triple-star system HD 131399ABC. (A to D) The central regions that are affected by the coronagraph and residual scattered starlight are blocked by a mask (dashed circles), with the location of star A indicated by the crosshairs. (E) Composite of the point spread function (PSF)-subtracted region (dashed region) superposed on the wide-field *K1* image showing the stellar components of the system, whose

luminosities are adjusted to the level of the planet for clarity. In each image, the luminosity of component A (but not components B and C) has been suppressed by the use of a coronagraph. The images in panels (A) to (C) were processed with angular and spectral differential imaging to subtract the stellar PSF, whereas panel (D) and the PSF-subtracted region of (E) were processed only with angular differential imaging (10). Images in (A) to (D) share the same field orientation.

Table 1. Basic parameters of the stars and directly imaged planet in HD 131399. The mass, effective temperature, and spectral type of the previously unresolved B and C stars (except where noted) were estimated from their <i>K1</i> luminosity (17–19, 35). The planet’s temperature and spectral type were determined through spectral fitting (see next section on characterization). Apparent <i>J</i> , <i>H</i> , and <i>K</i> magnitudes for HD 131399A were obtained from (36). M_{\odot} , solar mass; N/A, not applicable.				
Parameter	HD 131399A	HD 131399Ab	HD 131399B	HD 131399C
Spectral type	A1V*	T2 to T4	G	K
Mass	$1.82 M_{\odot}^{\dagger}$	$4 \pm 1 M_{\text{Jup}}$	$0.96 M_{\odot}^{\dagger}$	$0.6 M_{\odot}$
Effective temperature (T_{eff})	9300 K	850 ± 50 K	5700 K	4400 K
Projected separation from A (arc sec)	N/A	0.839 ± 0.004 (June 2015) 0.834 ± 0.004 (March 2016) 0.830 ± 0.004 (May 2016)	3.149 ± 0.006 (June 2015) 3.150 ± 0.006 (March 2016) 3.149 ± 0.006 (May 2016)	3.215 ± 0.006 (June 2015) 3.220 ± 0.006 (March 2016) 3.220 ± 0.006 (May 2016)
Position angle (degrees E of N from A)	N/A	194.2 ± 0.3 (June 2015) 193.8 ± 0.3 (March 2016) 193.5 ± 0.3 (May 2016)	221.9 ± 0.3 (June 2015) 221.5 ± 0.3 (March 2016) 221.8 ± 0.3 (May 2016)	222.0 ± 0.3 (June 2015) 221.9 ± 0.3 (March 2016) 222.1 ± 0.3 (May 2016)
<i>J</i> magnitude	6.772 ± 0.018	20.0 ± 0.2	N/A	N/A
<i>H</i> magnitude	6.708 ± 0.034	19.7 ± 0.2	N/A	N/A
<i>K</i> -band magnitude	$K = 6.643 \pm 0.026$	$K1 = 19.1 \pm 0.1$	$K1 = 8.5 \pm 0.1$	$K1 = 10.5 \pm 0.1$
* (37). † (38).				

Centaurus-Lupus association [UCL (11–13)] at a distance of 98 ± 7 pc (14) whose basic properties are given in Table 1. The system's membership in UCL is confirmed by its parallax and kinematics (11–13), and the well-constrained age of the association provides greater confidence in the young age of the system than for most directly imaged exoplanet host stars (see supplementary text for the detailed age analysis). Despite its youth, the system shows no evidence of infrared excess, and thus its primordial disk has probably been depleted to beneath detectable levels (15).

We observed HD 131399 on 12 June 2015, obtaining a wide range of near-infrared spectral coverage ranging from the *Y* band to the *K* band (0.95 to 2.25 μm) and diffraction-limited imaging with an 8.2-m telescope aperture. Our observations (10) resulted in the discovery of HD 131399Ab, a point source with a 10^{-5} contrast to HD 131399A and a projected separation of 0.84 arc sec, or 82 ± 6 AU (Fig. 1 and Table 1). After the initial discovery, we obtained follow-up observations (10) to verify whether the faint source is physically associated with the parent star (i.e., shares common proper motion) and to improve the quality of the near-infrared spectrum, enabling characterization of the planet's atmospheric properties.

We detected HD 131399Ab with a signal-to-noise ratio in the *Y* (1.04 μm), *J* (1.25 μm), *H* (1.62 μm), *K1* (2.11 μm), and *K2* (2.25 μm) bands of 9.3, 13.2, 15.5, 23.5, and 11.9, respectively. Following astrometric calibrations (10), we measured a positional displacement to HD 131399A of $\Delta\alpha$ (right ascension) = 12 ± 8 milli-arc sec (mas) and $\Delta\delta$ (declination) = 6 ± 8 mas over the 11-month baseline, where the uncertainties are dominated by the calibration of the instrument orientation across the two epochs. This allows us to reject the hypothesis of a background object, which would have moved relative to HD 131399A by $\Delta\alpha = 27.3 \pm 0.6$ mas and $\Delta\delta = 28.8 \pm 0.6$ mas due to the relatively high proper motion of the system (14). Assuming a Keplerian orbit for the planet with a semimajor axis equivalent to its projected separation of 82 AU yields a period of ~ 550 years, which, for a face-on circular orbit over 11 months, is expected to produce ~ 9 mas of relative motion, consistent with our observations.

The bound planet hypothesis is also supported by the low probability of detecting an unbound object within UCL that happens to share a similar spectral type to HD 131399Ab (as discussed in the next section). Following the arguments in (16), the false-alarm rate of an unassociated object with a planetlike spectrum per field of view is $\sim 2 \times 10^{-7}$. The total false-alarm probability of one such object appearing in our 33 fields of view (so far explored in our survey) is given by the binomial distribution, resulting in a probability of $\sim 6.6 \times 10^{-6}$. Although the probability of detecting a bound giant planet is not yet well established, results from the first several hundred stars surveyed suggest that this value is around a few percent—orders of magnitude higher than

the probability of detecting an unbound object with a planetlike spectrum.

Characterization of HD 131399Ab

We convert the planet's *J*-, *H*-, and *K1*-band aperture photometry to a mass estimate via compar-

ison to widely used evolutionary tracks for hot-start initial conditions (16–18), in which the planet retains its initial entropy of formation. Systematic interpolation between hot-start evolutionary tracks yields a mass of 4 ± 1 Jupiter masses (M_{Jup}), which places HD 131399Ab firmly in the planetary mass

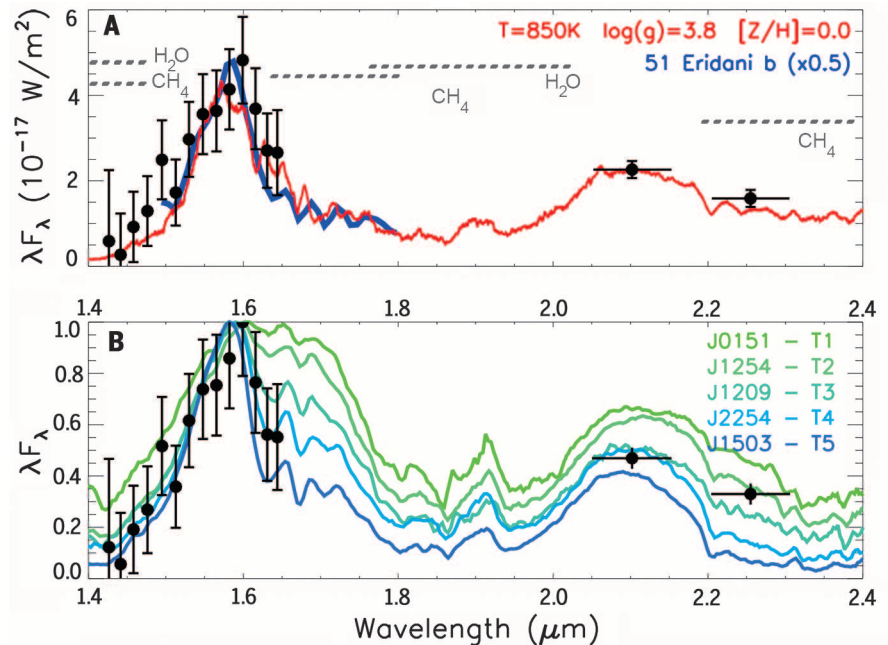


Fig. 2. Near-infrared spectrum of HD 131399Ab. (A) HD 131399Ab spectrum (black) alongside the best-fit model atmosphere in red (18), with $T_{\text{eff}} = 850$ K and $\log(g) = 3.8$ cm/s², showing water and methane absorption in the atmosphere with the approximate absorption regions indicated by the gray dashed lines. The spectrum of the T-type exoplanet 51 Eri b (16) is shown in blue, scaled by 50% to roughly match the luminosity of HD 131399Ab. F_λ , specific flux; Z/H , metallicity. (B) Near-infrared spectrum of HD 131399Ab and spectra of standard field brown dwarfs (39, 40), with each 1.4- to 2.4- μm spectrum normalized independently in λF_λ units (equivalent to power per unit area). The objects' labels correspond to the object designations from the Two Micron All Sky Survey (J2000 hours and minutes of right ascension) and the spectral type. Vertical error bars indicate 2σ photometric uncertainties; horizontal bars denote photometric bandpass.

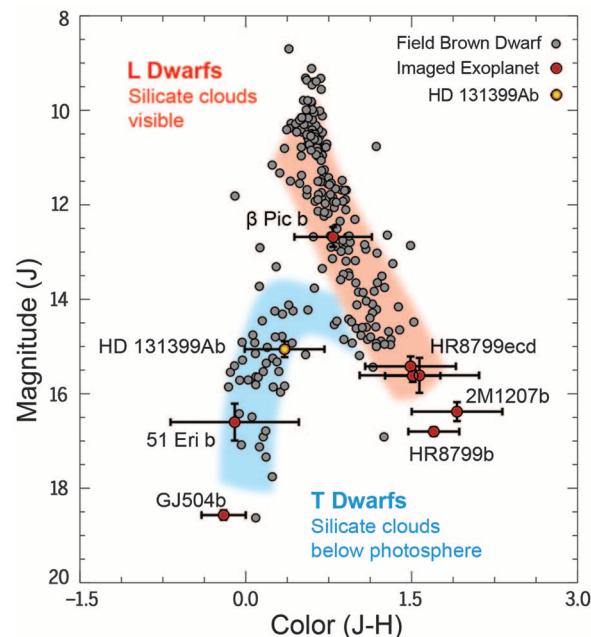


Fig. 3. J-H color-magnitude diagram of brown dwarfs and directly imaged giant exoplanets. HD 131399Ab falls among the methane-dominated T dwarfs near the L-T transition. The L- and T-dwarf data (with parallax-calibrated absolute magnitudes) were obtained from (41), whereas the directly imaged exoplanet data are from (16, 42–47). Vertical and horizontal error bars indicate 2σ photometric uncertainties.

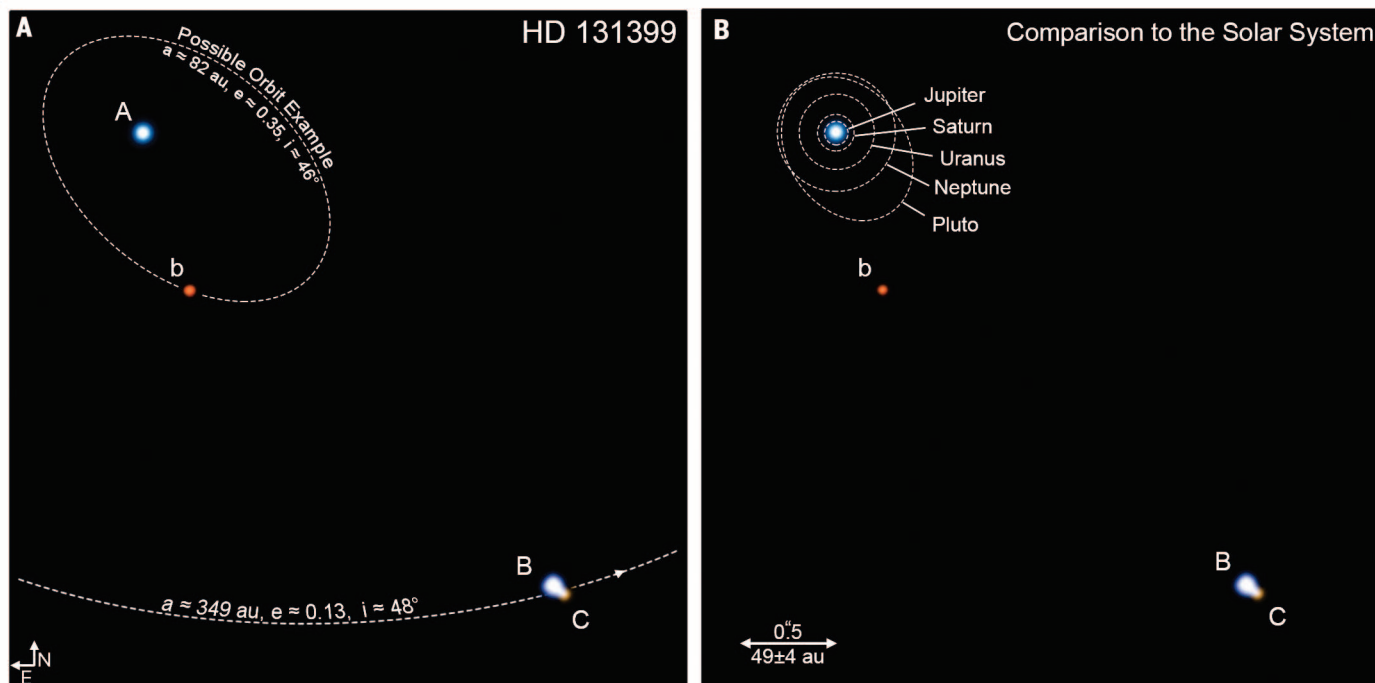


Fig. 4. Schematic illustration of the components of the HD 131399 hierarchical triple-star system and comparison to the solar system. (A) The dashed ellipse (top left) shows a preliminary orbit for the planet; the dashed curve at the bottom of this panel indicates our best-fit orbit of the BC pair. The orbit shown for the planet has orbital elements that are consistent with the data, although the astrometric uncertainties permit a substantial range of orbits (see the supplementary text for parameter ranges). (B) The image in (A) is reproduced with the

orbits of the solar system planets overlaid. The underlying image is a composite image of the actual PSFs superposed on a dark-sky background. The image is composed of SPHERE *J*-, *H*-, and *K*-band PSFs for components A and Ab (colored as blue, green, and red, respectively) and the monochromatic *K*-band PSF of components B and C. For clarity, the luminosity of the planet is enhanced by a factor of 10^5 , and because only *K*-band photometry exists for components B and C, their colors have been adjusted to be representative of typical G and K stars.

regime. Even in the unlikely event that the system is much older (by a few hundred million years), companion Ab would necessarily be of planetary mass ($<13 M_{\text{Jup}}$). The *H* versus *H-K* color of the planet is inconsistent with the cold-start scenario, in which the planet has lost some fraction of its initial entropy due to inefficient accretion (19, 20), but consistent with hot-start models including a partly cloudy atmosphere and/or super-solar metallicity (fig. S1).

Using the integral field spectrograph (2I) on SPHERE, we obtained a 0.95- to 1.65- μm spectrum. This spectrum allows the characterization of water and methane absorption bands within 1.4 to 1.6 μm , although the signal-to-noise ratio in the individual spectral channels at shorter wavelengths is too poor to be useful in spectral analysis. In the *K*-band, where the contrast with the star is more favorable, the dual-band images also probe the 2.2- μm methane absorption. Like the exoplanet 51 Eridani b (16) and other field (non-exoplanet) T-type brown dwarfs, the near-infrared spectrum of HD 131399Ab (Fig. 2) displays prominent methane and water absorption bands. The data are in good agreement with models of exoplanetary atmospheres (18), allowing us to estimate the atmospheric properties of effective temperature (T_{eff}) and surface gravity (g). Systematic exploration of interpolated atmospheric models indicates $T_{\text{eff}} = 850 \pm 50 \text{ K}$ and $\log(g) = 3.8^{+1.7}_{-0.8}$ (centimeters per square second), where the uncertainty in surface

gravity is mostly dominated by systematic uncertainties within the models (namely in the cloud properties) and not by the model-data fit. Comparison to standard classifications of field brown dwarfs (Fig. 2B) indicates a spectral type of T2 to T4.

The transition between L and T spectral types ($T_{\text{eff}} \sim 2100$ to 1300 K and $T_{\text{eff}} \sim 1300$ to 600 K, respectively) is marked by the appearance of *H*- and *K*-band methane absorption in the atmospheres of the cooler T dwarfs. In *J* versus *J-H* color-magnitude space (Fig. 3), this appears as bluer color (more negative *J-H* color) compared with the hotter L dwarfs. At the threshold of the L-T transition, the photosphere becomes brighter in the *J*-band as silicate clouds transition from above to below the photosphere (22). The fact that cloudy directly imaged exoplanets (such as HR8799bcde, β Pic b, or 2M1207b) appear at the bottom of the L-dwarf sequence points to cloud layers in these low-gravity objects that are thicker than in their higher-gravity brown dwarf counterparts (23, 24). In contrast, HD 131399Ab and the two other directly imaged T-type exoplanets follow the T-dwarf sequence, which we interpret as evidence for a similarity between the mostly or fully cloud-free atmospheres of these exoplanets and cool field brown dwarfs. HD 131399Ab is the closest directly imaged exoplanet to the L-T transition, which is consistent with the partly cloudy atmosphere suggested by

the *H* versus *H-K* hot-start model predictions (fig. S1).

Orbital characterization of HD 131399

HD 131399Ab is the widest known exoplanet that orbits within a triple system (Figs. 4 and 5). Because the presence of a second and third star can greatly limit the phase space where planetary orbits are stable, observing a system in this configuration is thought to be unlikely (7, 25). In our ongoing survey, we have imaged 18 single A-type stars and 15 binary or triple-star systems with separations similar to HD 131399A-BC. Although the sample size is small, it is surprising to us that the first planet detected in our survey is in a triple system.

We used astrometric observations dating back to 1897 (table S2) (26) to fit the orbit to a grid of models for a binary system using the center of mass of the BC system. Our neglect of the BC orbit is motivated by the system's hierarchical nature and the fact that most previous data could not resolve the pair. Our best-fit model (fig. S2 and table S3) consists of a semimajor axis of $a_* = 349 \pm 28 \text{ AU}$, eccentricity of $e_* = 0.13 \pm 0.05$, and inclination of $i_* = 45^\circ$ to 65° with respect to the plane of the sky, where the $*$ subscript denotes the values for the BC orbit around HD 131399A. Using only the newer, more reliable data permits a wider range of $a_* = 270$ to 390 AU, $e_* = 0.1$ to 0.3, and $i_* = 30^\circ$ to 70° . The available astrometry for the planet does not permit a

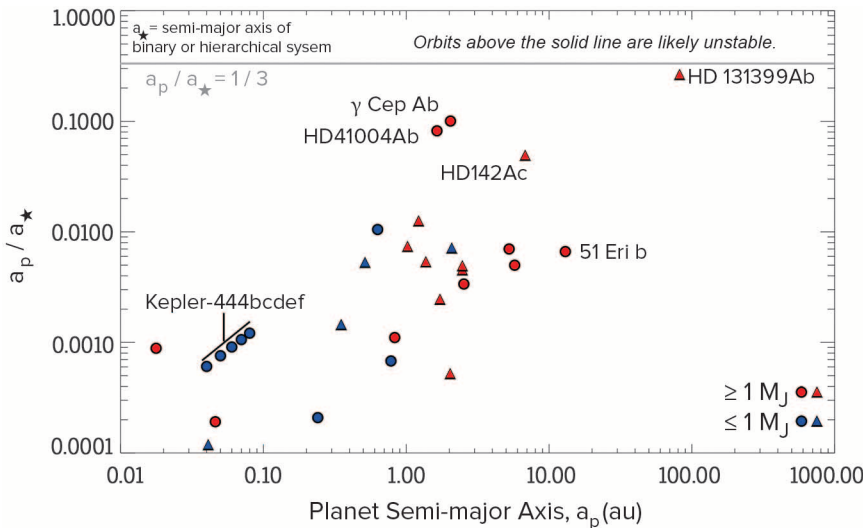


Fig. 5. Ratio of semimajor axes of planets that orbit one star of a multiple system (satellite, or S-type, planets) to the semimajor axes of their host systems. The gray solid line at one-third times the binary separation represents the approximate critical radius of tidal truncation and orbital stability in the coplanar case (25). Although the critical radius varies somewhat for different parameters of stellar mass ratio, eccentricity, and inclination, HD 131399Ab is much closer to the critical radius than any other known exoplanet. For systems in which either the planet or stars lack precise orbital solutions, their projected separations are plotted instead (denoted by triangular plot points instead of circles). This includes HD 131399Ab, although from the results of the preliminary orbit fit, the semimajor axes of this system are indeed similar to the projected separations. See table S4 for the list of included objects and their associated references. M_J , Jupiter mass.

robust orbital solution, though we performed a preliminary orbit fit to obtain the plausible parameter ranges of $a_{\text{planet}} = 82^{+23}_{-27}$ AU, $e_{\text{planet}} = 0.35 \pm 0.25$, and $i_{\text{planet}} = 40^{+80}_{-20}^\circ$, with no single solution being strongly preferred.

The orbital configuration of HD 131399 results in a more dynamically extreme configuration than for any known exoplanet within a binary or multiple system (Fig. 5 and table S4), with the ratio of semimajor axes $q = a_{\text{planet}}/a_* = 0.14$ to 0.38. Values of $q < 0.23$ require higher planetary eccentricities ($e_p > 0.3$) to maintain the ≥ 82 AU observational constraint on the planet's projected separation. The most dynamically similar planets to HD 131399Ab are γ Cephei Ab (27), discovered via radial velocity measurements; HD 41004Ab (28); and HD 142Ac (29), for which $q \sim 0.1$. Perhaps the most similar well-studied example is the transiting system Kepler-444, which hosts five sub-Earth-sized planets within 0.1 AU from the primary Kepler-444A (30). The latter stellar system is likewise a hierarchical triple, with a tight M-dwarf binary at 66 AU from the planet-hosting primary star. Though similar to these other systems, HD 131399 stands out due to the proximity of the planet's orbit to that of the other stars in the system.

We use a small suite (~ 300) of N-body simulations (10) to demonstrate that stable orbital configurations that are consistent with the astrometric constraints exist for all four bodies. This holds even for some of the more extreme configurations (i.e., smaller A-BC semimajor axis and higher eccentricity). The current astrometry also permits unstable orbits for the planet. Given the young age of the system, the planet might be on

an unstable orbit, perhaps due to planet-planet or planet-star scattering, and could yet be ejected to become a free-floating planetary-mass object. This is not the most likely scenario, as the time scale for the planet to suffer an ejection or collision is only a few million years (25). In all cases, the orbit of HD 131399Ab is non-Keplerian, as the planet's orbital parameters (a , e , and i) undergo complex evolution due to the influence of the BC pair (fig. S3).

Formation of HD 131399Ab and the origin of its long-period orbit

Given its location in a triple system, a broad set of formation pathways is possible for HD 131399Ab. Because planet formation is inhibited in the outer disk regions due to the strong perturbations from the binary (31, 32), it is unlikely that HD 131399Ab formed in isolation on its present long-period orbit around HD 131399A. We speculate that the planet may have arrived at its present orbit through one of three possible scenarios. Scenario (i): The planet formed on a short orbit around star A and subsequently underwent a planet-planet scattering event that ejected it to its current long-period orbit (33). This scenario requires the presence of a massive planet on a shorter-period orbit. Such a planet could have evaded detection if it were beneath our sensitivity limits (see supplementary online text for details). As a consequence we would also expect the Ab orbit to be rather eccentric. Scenario (ii): HD 131399Ab formed as a circumbinary planet around components B and C and underwent a scattering event via interactions with another planet or with the binary itself

(6). This scenario would also be most consistent with an eccentric Ab orbit. Scenario (iii): The planet formed around either component before the A-BC system arrived in its present configuration. The stellar orbits could have evolved subsequently due to interactions with the natal disks or secular effects (34). This scenario does not require the presence of a second close-in massive planet, though the resulting outer planetary orbit may be indistinguishable. Thus, it is possible that the planet is no longer orbiting the star around which it formed. These scenarios are also consistent with HD 131399Ab obtaining an orbit around all three components, although the short lifetime of such an orbit makes this configuration unlikely.

REFERENCES AND NOTES

- N. Batalha et al., *Astrophys. J. Suppl. Ser.* **204**, 24 (2013).
- J. Winn, D. Fabrycky, *Annu. Rev. Astron. Astrophys.* **53**, 409–447 (2015).
- B. P. Bowler, <http://adsabs.harvard.edu/abs/2016arXiv160502731B> (2016).
- G. Duchêne, A. Kraus, *Annu. Rev. Astron. Astrophys.* **51**, 269–310 (2013).
- D. Veras, J. Crepp, E. Ford, *Astrophys. J.* **696**, 1600–1611 (2009).
- R. A. Smullen, K. M. Kratter, A. Shannon, *Mon. Not. R. Astron. Soc.* **10.1093/mnras/stw1347** (2016).
- N. Moeckel, D. Veras, *Mon. Not. R. Astron. Soc.* **422**, 831–840 (2012).
- K. M. Kratter, H. B. Perets, *Astrophys. J.* **753**, 91 (2012).
- J.-L. Beuzit et al., *Proc. SPIE* **7014**, 701418 (2008).
- Materials and methods are available as supplementary materials on Science Online.
- P. T. de Zeeuw, R. Hoogerwerf, J. de Bruijne, A. Brown, A. Blaauw, *Astron. J.* **117**, 354–399 (1999).
- M. Pecaut, E. Mamajek, E. Bubar, *Astrophys. J.* **746**, 154 (2012).
- M. J. Pecaut, E. E. Mamajek, <https://arxiv.org/abs/1605.08789> (2016).
- F. van Leeuwen, *Astron. Astrophys.* **474**, 653–664 (2007).
- C. Chen, M. Pecaut, E. E. Mamajek, K. Y. L. Su, M. Bitner, *Astrophys. J.* **756**, 133 (2012).
- B. Macintosh et al., *Science* **350**, 64–67 (2015).
- I. Baraffe, G. Chabrier, T. Barman, F. Allard, P. H. Hauschildt, *Astron. Astrophys.* **402**, 701–712 (2003).
- F. Allard, D. Homeier, B. Freytag, *Philos. Trans. R. Soc. London Ser. A* **370**, 2765–2777 (2012).
- D. Spiegel, A. Burrows, *Astrophys. J.* **745**, 174 (2012).
- M. Marley, J. Fortney, O. Hubickyj, P. Bodenheimer, J. Lissauer, *Astrophys. J.* **655**, 541–549 (2007).
- R. Claudi et al., *Proc. SPIE* **7014**, 70143E (2008).
- A. Burrows, D. Sudarsky, I. Hubeny, *Astrophys. J.* **640**, 1063 (2006).
- T. Barman, B. Macintosh, Q. Konopacky, C. Marois, *Astrophys. J.* **733**, 65 (2011).
- M. Marley et al., *Astrophys. J.* **754**, 135 (2012).
- M. Holman, P. Wiegert, *Astron. J.* **117**, 621–628 (1999).
- D. Gill, *Astron. Nachr.* **144**, 89–94 (1897).
- A. Hatzes et al., *Astrophys. J.* **599**, 1383–1394 (2003).
- S. Zucker, T. Mazeh, N. Santos, S. Udry, M. Mayor, *Astron. Astrophys.* **426**, 695–698 (2004).
- A. Wittenmyer et al., *Astrophys. J.* **753**, 169 (2012).
- T. Dupuy et al., *Astrophys. J.* **817**, 80 (2016).
- R. Rafikov, K. Silsbee, *Astrophys. J.* **798**, 70 (2015).
- G. Lodato, F. Meru, C. J. Clarke, W. K. M. Rice, *Mon. Not. R. Astron. Soc.* **374**, 590–598 (2007).
- S. Chatterjee, E. Ford, S. Matsumura, F. Rasio, *Astrophys. J.* **686**, 580–602 (2008).
- L. G. Kiseleva, P. P. Eggleton, "The effect of tidal friction and quadrupolar distortion on orbits of stars or planets in hierarchical triple systems." *ASP Conf. Ser.* **154**, R. A. Donahue, J. A. Bookbinder, Eds. (Astronomical Society of the Pacific, 1998), p. 2118.
- I. Baraffe, G. Chabrier, F. Allard, P. H. Hauschildt, *Astron. Astrophys.* **337**, 403–412 (1998).
- M. Skrutskie et al., *Astron. J.* **131**, 1163–1183 (2006).
- N. Houk, *Michigan Catalogue of Two-dimensional Spectral Types for the HD stars. Volume 3. Declinations -40° to -26°* (University of Michigan, 1982).
- M. Kouwenhoven, A. Brown, S. F. Portegies Zwart, L. Kaper, *Astron. Astrophys.* **474**, 77–104 (2007).
- A. Burgasser et al., *Astron. J.* **127**, 2856–2870 (2004).

40. A. Burgasser, A. Burrows, J. D. Kirkpatrick, *Astrophys. J.* **639**, 1095–1113 (2006).
 41. T. Dupuy, M. Liu, *Astrophys. J. Suppl. Ser.* **201**, 19 (2012).
 42. A.-M. Lagrange et al., *Science* **329**, 57–59 (2010).
 43. M. Kuzuhara et al., *Astrophys. J.* **774**, 11 (2013).
 44. A. Zurlo et al., *Astron. Astrophys.* **587**, A57 (2016).
 45. T. Currie et al., *Astrophys. J.* **776**, 15 (2013).
 46. M. Janson et al., *Astrophys. J. Lett.* **778**, L4 (2013).
 47. S. Mohanty, R. Jayawardhana, N. Huélamo, E. Mamajek, *Astrophys. J.* **657**, 1064–1091 (2007).

ACKNOWLEDGMENTS

This work is based on observations performed with VLT/SPHERE under program IDs 095.C-0389A [principal investigator (PI): D.A.] and 296.C-5036A (PI: K.W.). K.W. is supported by the NSF Graduate Research Fellowship Program under grant 2015209499. The results reported herein benefited from collaborations and/or information exchange within NASA's Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by

NASA's Science Mission Directorate. This research has benefited from the SpeX Prism Spectral Libraries and the Washington Double Star Catalog maintained by the U.S. Naval Observatory at www.usno.navy.mil/USNO/astrometry/optical-IR-prod/wds/WDS. All atmospheric models used in this study can be found online at <http://svo2.cab.inta-csic.es/theory/newov/>, and all of the raw data products for HD 131399Ab and associated calibrations may be obtained from the ESO archive at <http://archive.eso.org/cms/eso-data.html>. SPHERE is an instrument designed and built by a consortium consisting of IPAG (Grenoble, France), Max-Planck-Institut für Astronomie (MPIA) (Heidelberg, Germany), Laboratoire d'Astrophysique de Marseille (Marseille, France), Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique (Paris, France), Laboratoire Lagrange (Nice, France), Istituto Nazionale di Astrofisica (INAF) Osservatorio di Padova (Italy), Observatoire de Genève (Switzerland), ETH Zurich (Switzerland), Nederlandse Onderzoeksschool voor de Astronomie (NOVA) (Netherlands), Office National d'Etudes et de Recherches Aérospatiales (France), and the Netherlands Institute for Radio Astronomy (ASTRON) (Netherlands) in collaboration with ESO. SPHERE was funded by ESO, with additional

contributions from CNRS (France), MPIA (Germany), INAF (Italy), FINES (Switzerland), and NOVA (Netherlands). SPHERE also received funding from the European Commission Sixth and Seventh Framework Programmes as part of the Optical Infrared Coordination Network for Astronomy (OPTICON) under grants R13-Ct-2004-001566 for FP6 (2004–2008), 226604 for FP7 (2009–2012), and 312430 for FP7 (2013–2016).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/673/suppl/DC1
 Materials and Methods
 Supplementary Text
 Figs. S1 to S4
 Tables S1 to S4
 References (48–81)

27 April 2016; accepted 24 June 2016
 Published online 7 July 2016
 10.1126/science.aaf9671

REPORTS

ORGANOMETALLICS

Isolation and structural and electronic characterization of salts of the decamethylferrocene dication

M. Malischewski,^{1*} M. Adelhardt,² J. Sutter,² K. Meyer,^{2*} K. Seppelt¹

Ferrocene and its decamethyl derivative [Cp*₂Fe] are the most common standards for nonaqueous electrochemical investigations because of their well-defined and only mildly solvent-dependent reversible Fe(II)/Fe(III) redox couple. Higher oxidation states have only rarely been studied. We report the isolation and crystallographic and spectroscopic characterization of surprisingly stable Fe(IV) salts of the [Cp*₂Fe]²⁺ dication, produced by oxidation of [Cp*₂Fe] with AsF₅, SbF₅, or ReF₆ in neat sulfur dioxide as well as [XeF](Sb₂F₁₁) in neat hydrogen fluoride. The Sb₂F₁₁[−] salt exhibits a metallocene with the expected mutually parallel arrangements of the Cp* rings, whereas the As₂F₁₁[−], AsF₆[−], SbF₆[−], and ReF₆[−] salts manifest tilt angles ranging from 4° to 17°. Both ⁵⁷Fe Mössbauer spectroscopy and superconducting quantum interference device magnetization studies reveal identical d-orbital splitting with an S = 1, ³E ground state based on the 3d electronic configuration e_{2g}³a_{1g}¹ of all [Cp*₂Fe]²⁺ salts.

Metallocenes, complexes with π -interactions between the central transition metal ion and two coplanar cyclopentadienyl (Cp = C₅H₅[−]) ligands, are prototypical compounds of historical importance for the field of organometallic chemistry. The iron analog [Cp₂Fe] was discovered in 1951 (1, 2), and the structure and bonding relationship of this ferrocene molecule and its derivatives have received great scientific attention ever since (3). Metallocenes and

their derivatives have now become integral parts of organometallic chemistry textbooks, classroom teaching, and laboratory classes. The ease of functionalization and the high stability of ferrocene under most reaction conditions led to numerous advances in the fields of medicinal organometallic chemistry and chemical catalysis, including commercial asymmetric and redox-switchable catalytic processes as well as carbonylation, hydrogenation, and polymerization (4).

The most stable oxidation states of the iron center in ferrocene are +2 and +3 (i.e., ferrocenium, [Cp₂Fe]⁺). These oxidation states are also very common for simple iron compounds, such as the binary fluorides FeF₂ and FeF₃, or [FeF₆]^{3/4−}, for which the concept of formal oxidation state assignment is most unambiguous. Literature reports of such simple complexes of iron in the +4

oxidation state are scarce, however. A rare example is the homoleptic tetrafluorido species [FeF₄], which was identified only by matrix isolation spectroscopy under cryogenic conditions (5). Simple compounds of higher formal oxidation state are usually stabilized by multiple-bonded ligands, as seen in the well-known Fe(VI) ferrate anion, [FeO₄]^{2−}, with terminal oxido ligands. Naturally occurring iron coordination complexes in the +4 oxidation state operate as metalloenzyme intermediates and have successfully been characterized on the basis of a number of model complexes. In all model complexes, this unusual oxidation state is stabilized by strong π -donor ligands, such as oxido (O^{2−}) or nitrido (N^{3−}) ligands (6–8). A small number of these high valent and generally very reactive iron complexes have even been structurally characterized, such as the first Fe(IV) oxo (9), Fe(IV) (10), and even an Fe(V) nitrido (11) complex. However, most iron complexes in unusually high oxidation states (i.e., > +4) are unstable, such as the Fe(V) oxo (12) and Fe(VI) nitrido (13) that were identified spectroscopically merely as fleeting intermediates. Purely organometallic compounds with an iron center in the +4 oxidation state are exceedingly rare. One example of an organometallic Fe(IV) complex is the highly reactive, thermally unstable (half-life of 30 hours at 23°C), tetrakis(1-norbornyl) iron complex [Fe(Nor)₄] (14, 15). The studies presented here, however, suggest that the classic cyclopentadienyl π -donor ligand pentamethylcyclopentadienide (Cp* = C₅Me₅[−], where Me = methyl) generally holds the potential to stabilize high oxidation states in organometallic chemistry, similar to the strong oxido and nitrido π -donor ligands in inorganic coordination chemistry (16).

The first observation that decamethylferrocenium [Cp*₂Fe]⁺ may be further oxidizable dates back to a 1980 report that an AlCl₃-based melt of decamethylferrocene exhibits two reversible electrochemical oxidation events (17). The electrochemical oxidation of the unsubstituted ferrocenium cation proved to be irreversible. In 1983, the dicationic decamethylferrocenium species was generated in liquid sulfur dioxide by means of coulometry and was characterized in situ by

¹Inorganic Chemistry, Institute of Chemistry and Biochemistry, Free University Berlin, 14195 Berlin, Germany.

²Inorganic Chemistry, Department of Chemistry and Pharmacy, Friedrich-Alexander-University Erlangen-Nürnberg, 91058 Erlangen, Germany.

*Corresponding author. Email: moritz.malischewski@fu-berlin.de (M.M.); karsten.meyer@fau.de (K.M.)

visible absorption spectroscopy and magnetic susceptibility measurements (18). The separation of the oxidized species from the electrolyte, however, was not achieved, and the oxidation of ferrocene with unsubstituted Cp ligands resulted in rapidly decomposing products, analogous to prior results (17, 18). Similarly, past attempts at chemical oxidation of ferrocene with very strong oxidizing agents (e.g., WF_6 , MoF_6 , and UF_6) have not met with success (19).

We report the successful two-electron chemical oxidation of decamethylferrocene, yielding the $[\text{Cp}^*_2\text{Fe}]^{2+}$ dication in isolable form as a salt with various counteranions. The stable salts enable x-ray crystallographic and spectroscopic characterization. The oxidation was achieved at room temperature by applying the powerful oxidizers AsF_5 , SbF_5 , and ReF_6 in liquid SO_2 , or $[\text{XeF}](\text{Sb}_2\text{F}_{11})$ with an excess of SbF_5 in HF solution (Fig. 1). The orange decamethylferrocene complex (18-electron species) oxidizes stepwise to yield the green ferric complex $[\text{Cp}^*_2\text{Fe}]^+$ (17-electron species) and finally the brown Fe(IV) species $[\text{Cp}^*_2\text{Fe}]^{2+}$ (16-electron complex). X-ray diffraction analyses of the $[\text{Cp}^*_2\text{Fe}]^{2+}$ salts with AsF_6^- , $\text{As}_2\text{F}_{11}^-$, SbF_6^- , $\text{Sb}_2\text{F}_{11}^-$, and ReF_6^- counteranions were performed, and the SbF_6^- and $\text{Sb}_2\text{F}_{11}^-$ salts were further characterized by temperature-dependent magnetic susceptibility measurements [superconducting quantum interference device (SQUID)] and ^{57}Fe Mössbauer spectroscopy. All isolated salts are stable at room temperature but must be handled and stored strictly under inert gas conditions. Although they are highly soluble in SO_2 or anhydrous HF, the use of organic solvents or exposure to air leads to immediate decomposition (20).

Strikingly, the determined crystal structures of the different $[\text{Cp}^*_2\text{Fe}]^{2+}$ salts (Fig. 2A) vary depending on the weakly coordinating counteranion. When crystallizing in the presence of a number of these counteranions, the Cp^* rings of certain dications are increasingly tilted: $\text{Sb}_2\text{F}_{11}^-$ (0°), $\text{As}_2\text{F}_{11}^-$ (4.17°), AsF_6^- (14.33°), ReF_6^- (16.17°), and SbF_6^- (16.56°) anions (fig. S10 and table S1). The larger anions such as $\text{M}_2\text{F}_{11}^-$ (M = metal) enforce a linear or nearly linear structure, whereas the smaller MF_6^- salts result in bent structures. In the molecular structure of the $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ salt, the Fe atom resides at the center of symmetry, leading to a parallel ring structure. The main characteristics of all crystal structures are intermolecular hydrogen bridges between the Cp^* methyl groups and fluorine atoms (distances >2.4 Å) and large distances between fluorine and iron atoms (>3.5 Å for MF_6^- and >4.6 Å for $\text{M}_2\text{F}_{11}^-$). Whereas the F-H structural features may induce or support the observed ring tilting, the long Fe-F distances suggest no bonding interactions between the ion pair, which we further confirmed by Hirshfeld electrostatic surface analysis (Fig. 2B) (21–23).

The bending of a metallocene, as observed in the molecular structures here, is generally seen only upon direct ligand coordination to the central metal ion (24). Bending without additional ligand coordination to the metal ion, although rare, is found in sterically strained ansa-metallocenes (25),

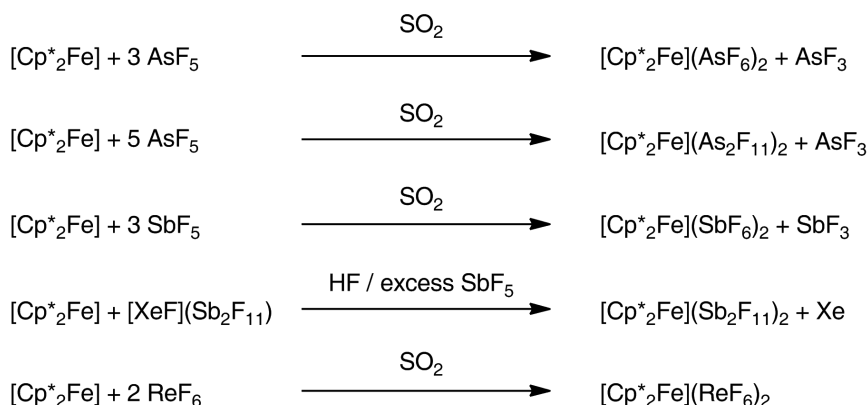


Fig. 1. Synthetic routes to the various $[\text{Cp}^*_2\text{Fe}]^{2+}$ salts.

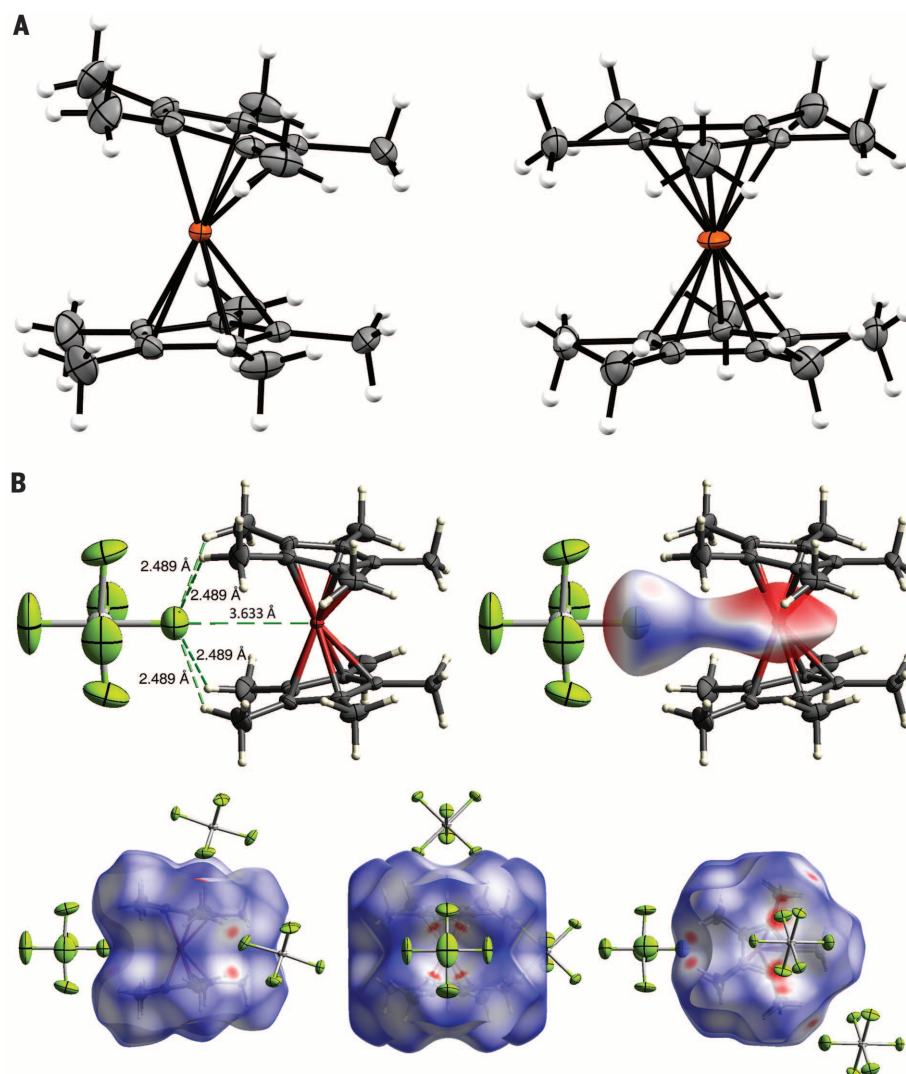


Fig. 2. Crystallographic analysis. (A) Molecular structures of $[\text{Cp}^*_2\text{Fe}]^{2+}$ in crystals of $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2 \cdot 2\text{HF}$ (left) and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ (right). Orange, Fe; gray, C; white, H; counteranions and cocrystallized HF are omitted for clarity (50% probability ellipsoid). (B) Representative distances between $[\text{Cp}^*_2\text{Fe}]^{2+}$ and SbF_6^- in crystals of $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2 \cdot 2\text{HF}$ (top left) with Hirshfeld surface analysis, highlighting regions of increasing intermolecular interactions from blue to red.

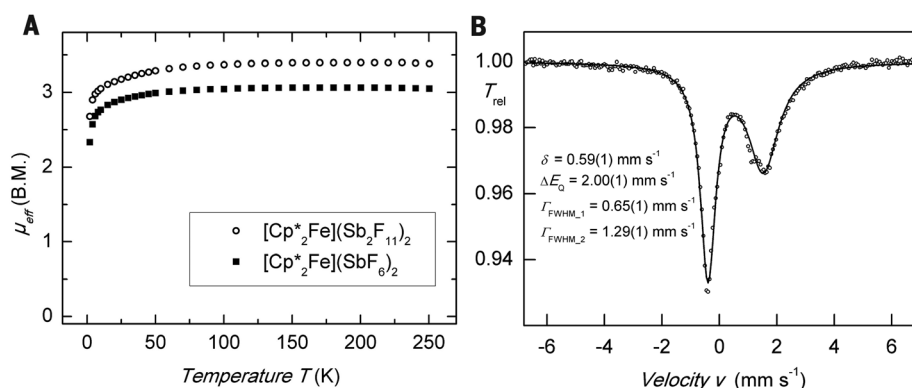


Fig. 3. Magnetochemical characterization. (A) SQUID magnetization data of $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2$ and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ recorded with an applied field of 1 T. (B) Zero-field ^{57}Fe Mössbauer spectrum of $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ in the solid state at 77 K. Circles represent the experimental data; black line depicts the best fit to the experimental data with parameters listed next to the spectrum. δ , isomer shift; ΔE_Q , quadrupole splitting; Γ_{FWHM} , line width at half maximum of peak height.

molecules with strong agostic interactions such as $[\text{Cp}^*_2\text{Ti}]^+$ (26), main-group metallocenes with sterically active free electron pairs (27), or heavy alkaline earth (27) and rare-earth metallocenes (28), for which the observed bending has been explained with the “polarizable ion model” (PIM) (29). The bent structures of $[\text{Cp}^*_2\text{Fe}](\text{MF}_6)_2$ may be explained by using the PIM as well: In the dicationic structures with the smaller MF_6^- counteranions, an induced dipole moment of the dication leads to a stronger electrostatic anion-cation interaction and resulting tilting. The observed tilting is further enforced by F-H interactions between the fluorinated anion and the Cp^* methyl groups. Because of the weaker basicity and the resulting weaker interaction of the larger, more diffusely charged, $\text{M}_2\text{F}_{11}^-$ counterions with $[\text{Cp}^*_2\text{Fe}]^{2+}$, the Cp^* tilting in the crystal structure is minimal relative to the MF_6^- salts (30).

DFT calculations (Gaussian09, M06L/def2-TZVPP) revealed that the energy difference between the parallel and bent structure of $[\text{Cp}^*_2\text{Fe}]^{2+}$ is exceedingly small. For example, the energy difference between an optimized D_{5d} symmetric (staggered conformation) parallel structure and an optimized bent structure (C_{2v}) with eclipsed methyl groups and tilt angles ranging from 3.11° to 16.77° is only 1.8 and 12.9 kJ/mol, respectively (table S2). Packing effects easily compensate for this small energy difference. The calculated conformational effects are also in accordance with the experimental data: All bent structures of $[\text{Cp}^*_2\text{Fe}]^{2+}$ show an eclipsed conformation of the Cp^* rings, whereas the parallel structure is staggered. In all cases, the methyl groups are slightly bent away from the Cp^* plane. The experimental and calculated $\text{Cp}^*(\text{centroid})\text{-Fe}$ distances are all found between 1.75 and 1.76 Å. The experimentally observed Fe-C distances are in the range of 2.12 to 2.15 Å, whereas the calculated values have a higher variability (2.076 to 2.178 Å). The Fe-C bond lengths found here are significantly longer than those reported for $[\text{Cp}^*_2\text{Fe}](\text{PF}_6)_2$ (2.085 ± 0.003 to 2.128 ± 0.002 Å) or $[\text{Cp}^*_2\text{Fe}]$ (2.045 ± 0.003 to 2.053 ± 0.002 Å)

(31, 32). This elongation of the Fe-C bonds with increasing oxidation state seems counterintuitive but is a consequence of the stepwise electron removal from weakly bonding orbitals of the ferrous starting material. The C-C bond lengths of the Cp^* rings are all in the expected range and are similar to those of reported metallocenes. The structure of $[\text{Cp}^*_2\text{Fe}]^{2+}$ ($3d^4$) is notably more flexible, with tilt angles ranging from 4° to 17° , which is in surprising contrast to the strictly parallel ring structures of $[\text{Cp}^*_2\text{Fe}]$ ($3d^6$) and $[\text{Cp}^*_2\text{Fe}]^+$ ($3d^5$).

^{57}Fe -enriched samples of $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2$ and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ were prepared and carefully purified by crystallization for investigation by ^{57}Fe Mössbauer spectroscopy and SQUID magnetization measurements performed on identical samples (Fig. 3). All data were checked for reproducibility by repeated measurements of various independently synthesized and purified samples (figs. S2 to S4). At 250 K, the samples $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2$ and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ exhibit effective magnetic moments, μ_{eff} , of 3.04 and $3.38 \mu_B$ (where μ_B is the Bohr magneton). The measurements show negligible temperature dependence in the range 50 to 250 K. Below 50 K, the magnetic moment gradually decreases because of zero-field splitting and finally reaches values of 2.33 and $2.68 \mu_B$ at 2 K, respectively. Although the room-temperature moments are higher than the calculated spin-only value of $2.83 \mu_B$, these values clearly specify an $S = 1$ ground state for the Fe(IV) species. This is in accordance with previously reported susceptibilities of in situ generated Fe(IV) dications (18) and is in excellent agreement with μ_{eff} values reported for the isoelectronic $[\text{Cp}_2\text{Cr}]$ ($3.20 \pm 0.16 \mu_B$) (33) with an orbitally degenerate $^3E_{2g}$, $(e_{2g})^3(a_{1g})^1$ ground state (34). We therefore also suggest a 3E ground state for the newly synthesized $[\text{Cp}^*_2\text{Fe}]^{2+}$ with notable contributions of first-order spin-orbit coupling, resulting in $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2$ and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ room-temperature magnetic moments that deviate from the spin-only values.

To further substantiate the complexes' electronic configuration, we recorded zero-field ^{57}Fe

Mössbauer spectra of solid, ^{57}Fe -enriched samples in the temperature range 5 to 293 K (Fig. 3B and fig. S2). The Mössbauer spectra of $[\text{Cp}^*_2\text{Fe}](\text{SbF}_6)_2$ and $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ at 77 K are very similar to each other, featuring asymmetric quadrupole doublets at $\delta = 0.59 \pm 0.02 \text{ mm s}^{-1}$ with electric quadrupole splittings $\Delta E_Q = 1.89 \pm 0.01 \text{ mm s}^{-1}$ (SbF_6^-) and $\Delta E_Q = 2.00 \pm 0.01 \text{ mm s}^{-1}$ ($\text{Sb}_2\text{F}_{11}^-$) (Fig. 3 and fig. S2). At 5 K, the spectrum of $[\text{Cp}^*_2\text{Fe}](\text{Sb}_2\text{F}_{11})_2$ displays a symmetrical quadrupole doublet at $\delta = 0.61 \pm 0.01 \text{ mm s}^{-1}$ with $\Delta E_Q = 2.01 \pm 0.01 \text{ mm s}^{-1}$ (fig. S2). With increasing temperature, the signal changes to an increasingly asymmetric form with respect to the line width (at 77 K) as well as intensity (at 293 K) (fig. S2). Summarizing the results of the experimental Mössbauer studies, the fluoroantimonate salts of linear and bent $[\text{Cp}^*_2\text{Fe}]^{2+}$ dications all show quadrupole doublets centered at $\sim 0.59 \text{ mm s}^{-1}$.

All observed isomer shifts of the tetravalent Fe ion in $[\text{Cp}^*_2\text{Fe}]^{2+}$ are slightly higher than those of the reduced ferric and ferrous species, which typically range between 0.48 and 0.50 mm s^{-1} at low temperatures (35–37). This trend in isomer shift correlates well with the elongated Fe- Cp^* distances found for the reported compounds. The longer and less covalent metal-ligand bonds result in less Fe 4s orbital participation and radial expansion, thereby effecting lower core s-electron density at the nucleus and, consequently, higher isomer shifts (38).

The large discrepancy between ΔE_Q values of the trivalent $[\text{Cp}^*_2\text{Fe}]^+$ salts ($\Delta E_Q = 0 \text{ mm s}^{-1}$) and tetravalent $[\text{Cp}^*_2\text{Fe}]^{2+}$ salts ($\Delta E_Q = 1.89$ to $2.01 \pm 0.01 \text{ mm s}^{-1}$) is expected for the metal-centered redox pair and originates from different valence contributions to the electric field gradient (EFG). The EFG value is estimated by summation of the incremental EFG expectation values of the corresponding valence electrons in their respective d-orbital ordering (Fig. 4). For approximate axial symmetry, the procedure can be restricted to the largest component ($V_{zz,\text{val}}$) along the unique axis of the EFG tensor, with expectation values of $\pm 4/7 e\langle r^{-3} \rangle$ for the (pure) dz^2 (a_{1g}) and dx^2-y^2 , dxy (e_{2g}) orbitals.

	[Cp* ₂ Fe](SbF ₆) ₂	[Cp* ₂ Fe](Sb ₂ F ₁₁) ₂
d(Fe-Cp*) (Å)	1.764	1.746
d(Fe-C) (Å)	2.142	2.131
∠(Cp*-Cp*) (°)	16.56	0
δ (mm s⁻¹)	0.59	0.59
ΔE_Q (mm s⁻¹)	1.89	2.00

Fe(II)	Fe(III)	Fe(IV)
<div><div>(-4/7)</div><div><div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div><div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div></div><div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div><div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div>↑</div><div>↓</div></div></div> <div><div><div>↑</div><div>↓</div></div><div><div></div></div></div>		

Fig. 4. Structural parameters and electronic structures. Relevant bond lengths and tilt angles (as defined in fig. S10), ⁵⁷Fe Mössbauer parameters (at 77 K), and predicted order and occupation of the iron valence orbitals in [Cp₂Fe]ⁿ⁺, with $n = 0, 1, 2$; also shown are valence contributions of the individual d-electron expectation values to the EFG, $(V_{zz})_{\text{val}}$, in units of $e(r^{-3})$ (given in parentheses) and total valence contribution to the EFG (bottom line).

The electronic structure of the Fe(III) ion in [Cp₂Fe]⁺ and its decamethyl derivative can be described with an $a_{1g}^2 e_g^3$ configuration (1 + 2 molecular orbital ordering) (39). According to this model, $(V_{zz})_{\text{val}}$ sums up to $+4/7 e(r^{-3})$, which implies substantial quadrupole splitting [~ 4 mm s⁻¹ within the ionic limit (38)]. Given the experimental single-line spectra observed for trivalent [Cp₂Fe]⁺ and [Cp*₂Fe]⁺ (40, 41), a strong covalence contribution must engage in the total EFG. Evidently, the amount of this contribution is approximately the same as the valence contribution but has the opposite sign, which compensates $(V_{zz})_{\text{val}}$ and thus results in $\Delta E_Q = 0$. It is reasonable to assume that this covalence contribution must be present throughout the entire redox series. Interestingly, the experimental quadrupole splittings $\Delta E_Q = 1.89$ and 2.01 mm s⁻¹ found for the [Cp*₂Fe(IV)]²⁺ salts resemble that of [Cp*₂Fe(II)] (~ 2.4 mm s⁻¹) (37). Both experimental observations are in accordance with our proposed ligand field splitting, in which the a_{1g} orbital is situated above e_g (Fig. 4)—a configuration generally accepted for divalent ferrocenes (42). Summation of the valence contributions yields identical EFG values $(V_{zz})_{\text{val}} = +8/7 e(r^{-3})$ for the complexes with Fe(IV) and Fe(II). Because these EFGs translate into excessively large quadrupole splittings, the negative covalence contributions reduce $(V_{zz})_{\text{val}}$ to yield more reasonable and experimentally observed quadrupole splittings.

Although our proposed $e_g^3 a_{1g}^1$ configuration (2 + 1) is based on the SQUID magnetization study, which suggests an orbitally degenerate ³E ground state, the Mössbauer quadrupole splitting also would fit to the alternative $a_{1g}^2 e_g^2$ (1 + 2) configuration. In that case, rendering $(V_{zz})_{\text{val}} = 0$, the strong negative covalence contribution to the EFG would entirely determine the quadrupole

splitting, which would be of similar magnitude but opposite sign. Unfortunately, the two alternatives cannot be distinguished by sign-insensitive, zero-field Mössbauer spectroscopy, and their discrimination awaits future applied-field and theoretical investigations.

Despite shortcomings, this simplified crystal field model clearly predicts the experimentally observed, larger quadrupole splittings for the tetravalent [Cp*₂Fe]²⁺ dications relative to trivalent ferrocenium complexes. The Mössbauer spectroscopic study and SQUID magnetization data lend strong confidence to the conclusion that the isolated stable salts reported here display iron centers in the +4 oxidation state with the orbitally degenerate ³E ground state, and an (e_g)(a_{1g}) orbital ordering similar to the one reported for the archetypal ferrocene.

The study of a tetravalent metallocene thus adds a high-valent Fe(IV) derivative to the series of ferrocene complexes. Whereas the archetypal ferrocene and ferrocenium complexes are all linear, the tetravalent derivative emerges in linear and unusual bent structures.

REFERENCES AND NOTES

- G. Wilkinson, M. Rosenblum, M. C. Whiting, R. B. Woodward, *J. Am. Chem. Soc.* **74**, 2125–2126 (1952).
- E. O. Fischer, W. Pfab, *Z. Naturforsch. B* **7**, 377–379 (1952).
- J. Ruiz et al., *J. Am. Chem. Soc.* **120**, 11693–11705 (1998).
- R. D. Adams, *J. Organomet. Chem.* **637–639**, 1 (2001).
- T. Schlöder, T. Vent-Schmidt, S. Riedel, *Angew. Chem. Int. Ed.* **51**, 12063–12067 (2012).
- J. T. Groves, *J. Inorg. Biochem.* **100**, 434–447 (2006).
- J. Hohenberger, K. Ray, K. Meyer, *Nat. Commun.* **3**, 720 (2012).
- A. R. McDonald, L. Que Jr., *Coord. Chem. Rev.* **257**, 414–428 (2013).
- J.-U. Rohde et al., *Science* **299**, 1037–1039 (2003).
- C. Vogel, F. W. Heinemann, J. Sutter, C. Anthon, K. Meyer, *Angew. Chem. Int. Ed.* **47**, 2681–2684 (2008).

- J. J. Scepaniak et al., *Science* **331**, 1049–1052 (2011).
- F. Tiago de Oliveira et al., *Science* **315**, 835–838 (2007).
- J. F. Berry et al., *Science* **312**, 1937–1941 (2006).
- B. K. Bower, H. G. Tennent, *J. Am. Chem. Soc.* **94**, 2512–2514 (1972).
- R. A. Lewis et al., *Inorg. Chem.* **52**, 8218–8227 (2013).
- D. S. Williams, M. H. Schofield, J. T. Anhaus, R. R. Schrock, *J. Am. Chem. Soc.* **112**, 6728–6729 (1990).
- R. J. Gale, P. Singh, R. Job, *J. Organomet. Chem.* **199**, C44–C46 (1980).
- P. R. Sharp, A. J. Bard, *Inorg. Chem.* **22**, 2689–2693 (1983).
- K. Moock, L. Turowsky, K. Seppelt, *J. Fluor. Chem.* **37**, 253–258 (1987).
- H. Ogino, H. Tobita, H. Habazaki, M. Shimoi, *J. Chem. Soc. Chem. Commun.* **1989**, 828–829 (1989).
- J. J. McKinnon, M. A. Spackman, A. S. Mitchell, *Acta Crystallogr. B* **60**, 627–668 (2004).
- M. A. Spackman, D. Jayatilaka, *CrystEngComm* **11**, 19–32 (2009).
- M. A. Spackman, J. J. McKinnon, *CrystEngComm* **4**, 378–392 (2002).
- J. C. Green, *Chem. Soc. Rev.* **27**, 263–272 (1998).
- P. J. Shapiro, *Coord. Chem. Rev.* **231**, 67–81 (2002).
- M. W. Bouwkamp et al., *J. Am. Chem. Soc.* **124**, 12956–12957 (2002).
- D. J. Burke, T. P. Hanusa, *Comm. Inorg. Chem.* **17**, 41–77 (1995).
- W. J. Evans, L. A. Hughes, T. P. Hanusa, *J. Am. Chem. Soc.* **106**, 4270–4272 (1984).
- N. J. Long, *Metallocenes—An Introduction to Sandwich Complexes* (Blackwell Science, 1998).
- I. Krossing, I. Raabe, *Angew. Chem. Int. Ed.* **43**, 2066–2090 (2004).
- A. Sánchez Perucha, M. Bolte, *Acta Crystallogr. E* **63**, m1703 (2007).
- D. P. Freyberg, J. L. Robbins, K. N. Raymond, J. C. Smart, *J. Am. Chem. Soc.* **101**, 892–897 (1979).
- F. Engelmann, *Z. Naturforsch. B* **8**, 775–776 (1953).
- E. König, R. Schnakig, B. Kanellakopoulos, R. Klenze, *Chem. Phys. Lett.* **50**, 439–441 (1977).
- V. Gama et al., *Eur. J. Inorg. Chem.* **2000**, 2101–2110 (2000).
- A. Vertes, Z. Klencsar, M. Gal, E. Kuzmann, *Fullerene Sci. Technol.* **5**, 97–109 (1997).
- R. H. Herber, *Inorg. Chim. Acta* **291**, 74–81 (1999).
- P. Güttlich, E. Bill, A. X. Trautwein, *Mössbauer Spectroscopy and Transition Metal Chemistry—Fundamentals and Applications* (Springer, 2011).

39. R. Prins, J. F. Reinders, *J. Am. Chem. Soc.* **91**, 4929–4931 (1969).
 40. J. S. Miller, J. H. Zhang, W. M. Reiff, *Inorg. Chem.* **26**, 600–608 (1987).
 41. W. M. Reiff, *J. Appl. Phys.* **63**, 2957–2961 (1988).
 42. T. A. Albright, J. K. Burdett, M.-H. Whangbo, *Orbital Interactions in Chemistry* (Wiley, 1985).

ACKNOWLEDGMENTS

Supported by Free University Berlin, Fonds der Chemischen Industrie, Deutsche Forschungsgemeinschaft (GRK 1582, Fluorine

as a Key Element), and Friedrich-Alexander-University of Erlangen-Nürnberg. We thank J. Telser and E. Bill for helpful discussions and E. Zolnhofer for synthesis and magnetochemical characterization of $[Cp_2Cr]$. Computing time was made available by High-Performance Computing at ZEDAT/Free University Berlin. All crystal structures have been deposited with the Cambridge Crystallographic Data Centre under accession numbers CCDC 1449790 $[C_{20}H_{30}Fe(As_2F_{11})_2 \cdot SO_2]$, CCDC 1449791 $[C_{20}H_{30}Fe(AsF_6)_2 \cdot 2HF]$, CCDC 1449792 $[C_{20}H_{30}F_{14}FeRe_2]$, CCDC 1449793 $[C_{20}H_{30}Fe(Sb_2F_{11})_2]$, and CCDC 1449794 $[C_{20}H_{30}Fe(SbF_6)_2 \cdot 2HF]$.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/678/suppl/DC1
 Materials and Methods
 Figs. S1 to S10
 Tables S1 and S2
 References (43–53)

7 March 2016; accepted 14 July 2016
 10.1126/science.aaf6362

SOFT ELECTRONICS

Highly stretchable, transparent ionic touch panel

Chong-Chan Kim,^{1,*} Hyun-Hee Lee,^{1,*} Kyu Hwan Oh,^{1,2} Jeong-Yun Sun^{1,2,†}

Because human-computer interactions are increasingly important, touch panels may require stretchability and biocompatibility in order to allow integration with the human body. However, most touch panels have been developed based on stiff and brittle electrodes. We demonstrate an ionic touch panel based on a polyacrylamide hydrogel containing lithium chloride salts. The panel is soft and stretchable, so it can sustain a large deformation. The panel can freely transmit light information because the hydrogel is transparent, with 98% transmittance for visible light. A surface-capacitive touch system was adopted to sense a touched position. The panel can be operated under more than 1000% areal strain without sacrificing its functionalities. Epidermal touch panel use on skin was demonstrated by writing words, playing a piano, and playing games.

Integrated touch panels provide an easy and intuitive interface for interacting with display devices. Panels have been developed with several types of sensing systems, including resistive (1, 2), capacitive (3, 4), surface acoustic wave (5), and infrared (6) touch. Resistive touch-sensing and capacitive touch-sensing have become common in electronic devices such as mobile phones, computers, ticketing machines, point-of-sale terminals, and information kiosks (6, 7).

Both resistive and capacitive touch-sensing require transparent conducting films (TCFs). Indium tin oxide (ITO) has been mostly used owing to its sufficiently low sheet resistance (<200 ohms per square) and transparency (8). However, because the next generation of touch panels requires stretchability and biocompatibility to allow integration with the human body, touch panels based on ITO face issues owing to their brittle nature. Alternatives such as conducting polymers (9, 10), carbon nanotubes (CNTs) (11, 12), graphene (13, 14), and metal nanowires (15, 16) have been investigated for their combination of stretchability along with transmittance for visible light. However, the sheet resistance of these materials sharply increased when they were stretched, and these materials showed fatigue failure when repeatedly stretched (17). Furthermore, the biocompatibility

of these alternative materials still must be demonstrated (18, 19).

Hydrogels are hydrophilic polymer networks swollen with large amounts of water. Hydrogels are soft like tissue and very stretchable (20). Many hydrogels are biocompatible, so they can be used for drug delivery (21), tissue replacement (22), and wound healing (23). Some hydrogels are transparent, allowing for 99% transmittance for the full range of visible light (24), so they can be used to transmit optical information. Because hydrogels contain large amounts of water, which can dissolve ions, they can serve as ionic conductors. Strain sensors, pressure sensors (25, 26), and actuators (24) have been created with hydrogels as an ionic conductor by stacking conductor/insulator/conductor layers. Thus, hydrogels can be a key ingredient for the next generation of touch panels because of their stretchability, biocompatibility, and transparency. We demonstrate an ionic touch panel using a polyacrylamide (PAAm) hydrogel containing lithium chloride (LiCl) salts.

Among the various types of touch-sensing systems, a surface-capacitive system was adopted for our ionic touch panel. In a surface-capacitive touch system, the same voltage is applied to all corners of the panel, which results in a uniform electrostatic field across the panel. When a conductor, such as a human finger, touches the panel, the touch point becomes grounded, and a potential difference is generated between the electrode and the touch point. The potential difference causes current to flow from the electrode through

the finger. The magnitude of the current is determined by the distance between the touch point and electrode. As the distance decreases, a larger current is induced (27). This system is limited to a single touch, but it is simple in terms of structure because the panel consists of only one conductor layer and only four electrodes, with one at each corner.

The architecture of a one-dimensional (1D) ionic touch strip is shown in Fig. 1A. A strip of PAAm hydrogel containing LiCl salts (2 M) was used as an ionic conductor. The strip was connected to platinum (Pt) electrodes on both sides, and the same phase ac voltage was applied through current meters A1 and A2 to both ends. When a finger touched the strip, a closed circuit was formed because the finger was grounded, which allowed current to flow from both ends of the strip to the touch point. The signals were detected by current meters at each corner. The corresponding circuit diagram of the ionic touch strip is shown in Fig. 1B. The strip is virtually divided into two resistive parts by the touch point. Each resistor was connected with a capacitor of an electrical double layer and a current meter in series; these two parts were connected in parallel. This parallel circuit was connected in series with a capacitor created by a finger. In the parallel circuit of Fig. 1B, the effect of the capacitor that was formed by the electrical double layer was negligible because of the large capacitance of the electrical double layer and high operating frequency. The position of the touch point can be represented by a normalized distance, α . The left end and the right end of the strip correspond to $\alpha = 0$ and $\alpha = 1$, respectively.

The current flowing in two resistors can be represented by the following equations

$$I_1 \approx \frac{R_2}{R_1 + R_2} I_t = (1 - \alpha) I_t \quad (1)$$

$$I_2 \approx \frac{R_1}{R_1 + R_2} I_t = \alpha I_t \quad (2)$$

where I_1 and I_2 are the touching currents measured from current meters A1 and A2, respectively, and $I_t = I_1 + I_2$. The derivations of Eqs. 1 and 2 are shown in the supplement. The current change depending on the touch is shown in Fig. 1C. Baseline currents on the order of micro amperes were detected. The baseline current is a leakage current that flows through the parasitic capacitor formed between the panel and the environments (28). When a finger touches the gel strip, additional currents are drawn from the

¹Department of Materials Science and Engineering, Seoul National University, Seoul 151-742, South Korea. ²Research Institute of Advanced Materials (RIAM), Seoul National University, Seoul 151-744, South Korea.

*These authors contributed equally to this work. †Corresponding author. Email: jysun@snu.ac.kr

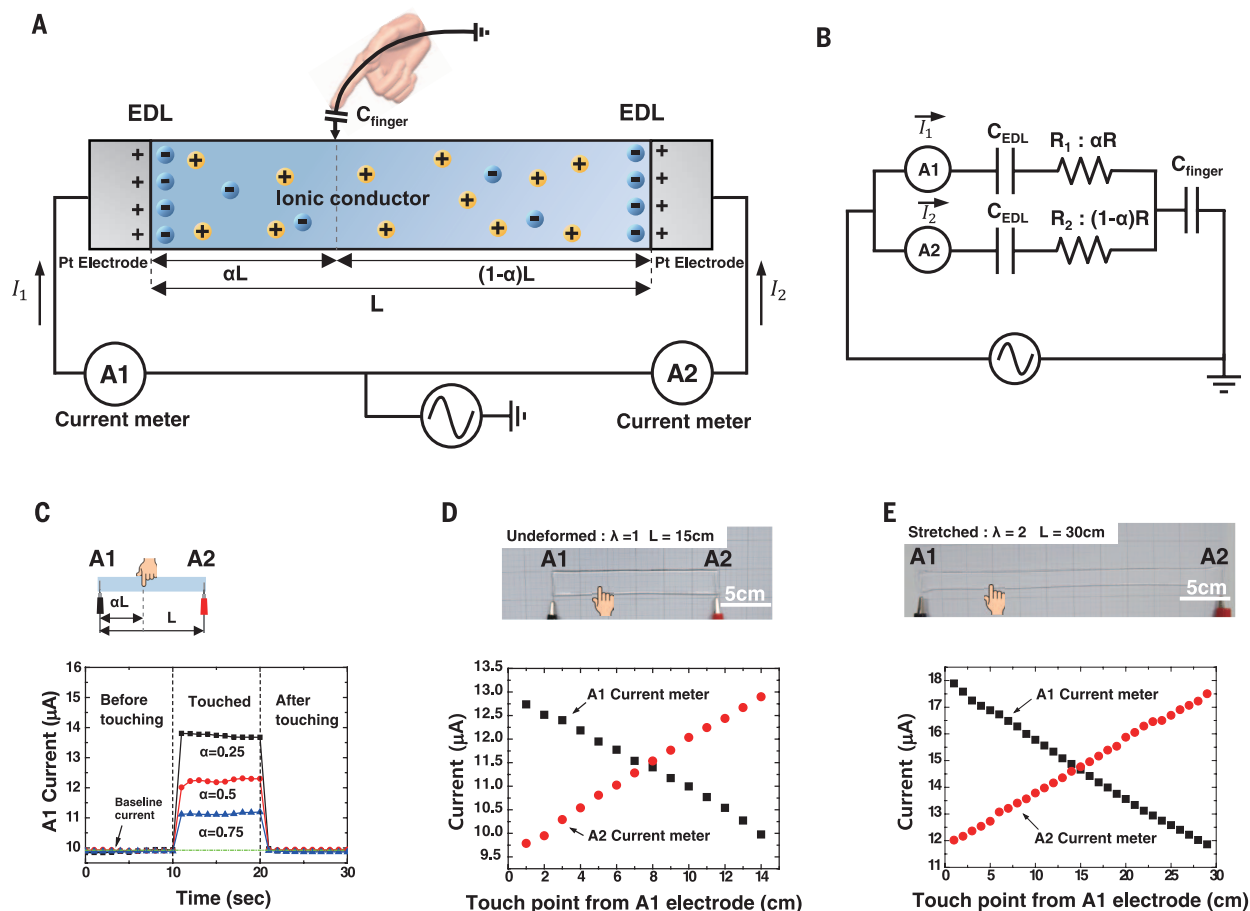


Fig. 1. A working principle of an ionic touch strip. (A) A schematic of a 1D ionic touch strip. When a finger touched the strip, a closed circuit was formed because the finger was grounded, which allowed current to flow from both ends of the strip to the touch point. (B) An electrical circuit diagram of the ionic touch strip. (C) The A1 current was recorded for various touch points ($\alpha = 0.25, 0.5$, and 0.75). Additional touching current flowed when the strip was touched by a finger.

The magnitude of the A1 current increased when the touch point became closer to the A1 electrode. The current returned to the baseline value when the strip was no longer being touched. (D) The average currents in a touched period were measured from A1 and A2 according to the distance from the touch point to the A1 electrode. (E) Linear relationship between the distance of the touch point and the current yield was maintained even after a stretching of $\lambda = 2$.

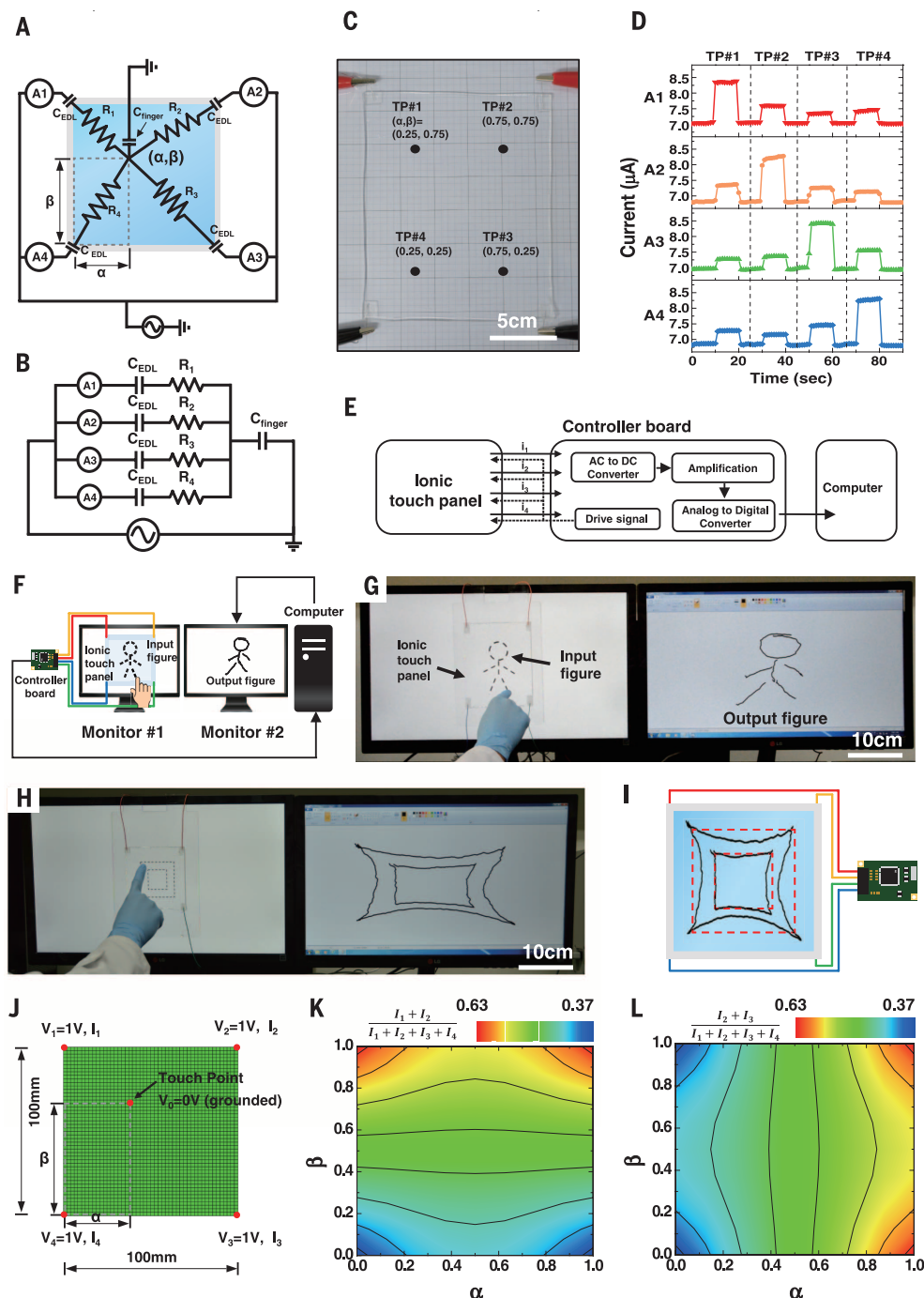
electrode to the finger (fig. S3). Here, we define this induced current as a “touching current.” The latency of the touching current was <20 ms (figs. S1 and S2). The touching current was proportional to the proximity of the electrode to the touch point. The strip was touched from the left electrode to the right electrode every 1 cm, and the measured currents from the A1 current meter and A2 current meter are shown in Fig. 1D. The sum of I_1 and I_2 remained constant, and I_1 linearly decreased, whereas I_2 increased as the touch point moved to the right. The resolution of the ionic touch strip will be limited by the resolution of the current meter. With a current meter, which has a resolution in the nanoampere range, the strip showed a resolution on the order of 10^{-4} m (fig. S6). The gel strip was stretched to two times its initial length [stretch (λ) = 2, length (L) = 30 cm]. When the gel strip was stretched (Fig. 1E), the parasitic capacitance of the strip increased because of the area expansion. Thus, the baseline current and the touching current both increased compared with that of the nonstretched states. In the stretched states,

the currents also showed similar negative linear correlations with the interval of the electrode and the touch points.

A 2D hydrogel panel was tested in order to detect the position of the touch point. As shown in Fig. 2, A and C, a thin hydrogel film, rectangular in shape, was connected to the Pt electrodes at each corner. Four current meters were also installed between the voltage source and each corner. Two normalized distances, α and β , were used to indicate the position of the touch on the panel. The bottom left corner of the panel corresponded to $(\alpha, \beta) = (0, 0)$, and the top right corner corresponded to $(\alpha, \beta) = (1, 1)$. When the panel is touched by a finger, it can be virtually divided into four resistive sections by the touch point (Fig. 2B). In the circuit, four virtual resistors are connected together in parallel, and this parallel circuit of resistors is connected to the capacitor by the finger in series. In Fig. 2C, the test positions of the touches are displayed. Touch points TP#1 (0.25, 0.75), TP#2 (0.75, 0.75), TP#3 (0.75, 0.25), and TP#4 (0.25, 0.25) were investigated. These points were

sequentially touched, and the current changes were measured from the four current meters installed at each corner. The measured currents from each current meter are plotted in Fig. 2D. When TP#1 was touched, the closest current meter, A1, from the touch point showed the largest value, and the farthest current meter, A3, displayed the lowest value. In this 2D case, the touching current gained from the four current meters was proportional to the proximity of the electrode to the touch point, which is similar to the case of the strip. The current values could be used to calculate the coordinate of the touch points in the 2D panels.

A controller board was designed to help communication between the ionic touch panel and a computer. A block diagram of the board is shown in Fig. 2E. The board generates a drive voltage signal, and the signal is applied to the corners of the panel. The board measures the currents to each corner. The currents were on the order of a few microamperes; thus, the currents were amplified and converted to digital form in order to calculate coordinates of the touch. Because the

Fig. 2. Position-sensing in a 2D ionic touch panel. (A) A schematic diagram of an ionic touch panel. A touched position was represented by two normalized distances, α and β . (B) An equivalent electrical circuit of the panel during touching. (C) The panel was fully transparent and stable during operation. To reveal the sensitivity of position detection, four touch points (TP#1 to TP#4) of the panel were investigated. (D) From TP#1 to TP#4, the points were sequentially touched, and the currents measured from the four current meters were plotted against time. (E) A block diagram of a controller board that helps communication between the ionic touch panel and a computer. (F and G) Operation of the ionic touch panel was demonstrated with a dashed-man drawing. Monitor #1 displays the dashed man, and an ionic touch panel was attached to monitor #1. When a finger drew through the trace of the dashed man on the panel, the drawing was detected and transferred to a computer by the controller board. The output drawing was displayed on monitor #2. (H and I) The ionic touch panel showed distortion near the edges. Two concentric squares were tested on the panel, and the outer square showed more distortion than that of the inner square. (J) A model for a FEM simulation. The same potential of 1 V was applied to the four corners of the panel, and the touch point at its position (α, β) was grounded as boundary conditions. (K and L) We calculated $\frac{I_1 + I_2}{I_1 + I_2 + I_3 + I_4}$ and $\frac{I_2 + I_3}{I_1 + I_2 + I_3 + I_4}$ from the simulation and plotted in (K) and (L), respectively. The nonlinearity of the resistance in a 2D panel caused the distortion shown in (H).

measured currents and coordinates showed negative linear correlation, the positions were calculated based on the following equations (27)

$$\alpha \propto \frac{I_2 + I_3}{I_1 + I_2 + I_3 + I_4} \quad (3)$$

$$\beta \propto \frac{I_1 + I_2}{I_1 + I_2 + I_3 + I_4} \quad (4)$$

where I_1, I_2, I_3 , and I_4 are increased currents by a touch measured at current meters A1, A2, A3, and A4, respectively. As shown in Fig. 2D, when a

finger touches the same α coordinate, the value of Eq. 3 is determined to be the same. The β coordinate shows the same tendency as the α coordinate.

The ionic touch panel was attached to monitor #1, which displayed an input figure (Fig. 2, F and G). A 1-mm-thick polymethyl methacrylate (PMMA) plate was applied between the panel and the monitor for an electrical insulation. The panel was connected to a computer through the controller board, and the computer displayed an output figure on monitor #2 based on a signal from the controller board. We drew a man on the

ionic touch panel (movie S1). Some distortion was observed on the edge of the output figures. Two concentric squares were drawn through the gel panel to test the distortion. As shown in Fig. 2, H and I, the outer square showed more distortion than the inner square. To investigate the reason for the distortion, we performed an electrical finite element method (FEM) simulation. The boundary conditions are displayed in Fig. 2J, and the electrical currents that flowed through the four corners were calculated at various touch points (α, β). A uniform electrical conductivity (1 S/m) was applied over

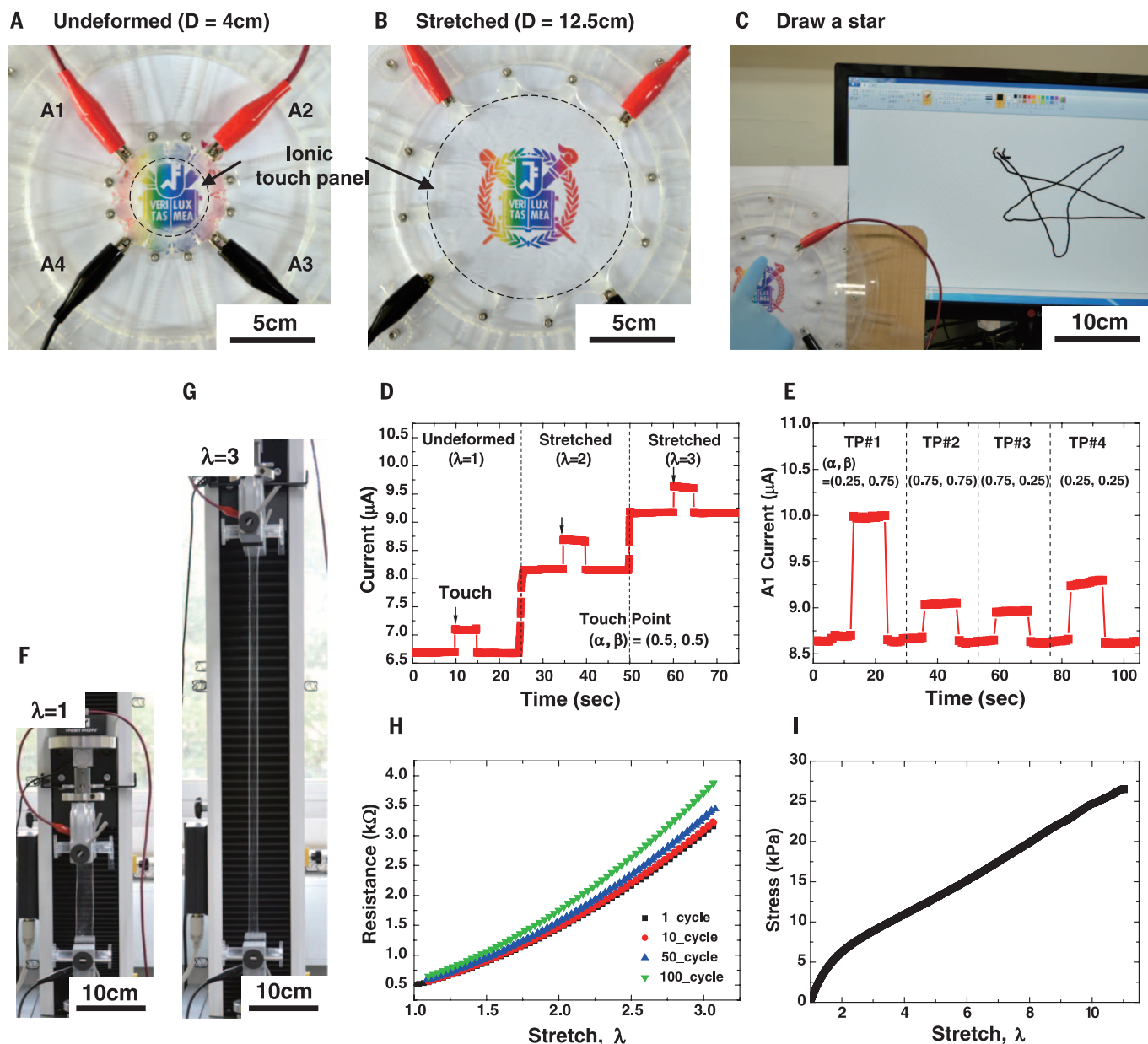


Fig. 3. A stretchable touch panel. (A and B) The ionic touch panel was stretched equibiaxially up to 1000% in area. A circular PAAm hydrogel film was glued to a biaxial stretcher, and the diameter of the hydrogel was increased from $D = 4$ cm to $D = 12.5$ cm. (C) The stretched touch panel was connected to a computer and operated as an input device. (D) The A1 current was measured before ($\lambda = 1$) and after stretching ($\lambda = 2$ and $\lambda = 3$). The baseline current increased according to the

stretch of the panel. However, the touching currents were insensitive to the stretching. (E) The A1 current was investigated for various touch points in a stretched state ($\lambda = 1.5$). (F and G) An electromechanical test was performed on a hydrogel containing LiCl salts (2 M) up to 100 cycles. (H) The resistance change of the gel strip was investigated during cyclic loadings. (I) A uniaxial tensile test was performed on a PAAm hydrogel.

the panel, and the potential of four corners was fixed at 1 V. The panel was split into 10 grids in rows and columns. Each intersection of the grids, which is regarded as a touch point, was sequentially grounded (0 V). Plotted in Fig. 2, K and L, respectively, are $(I_1 + I_2)/(I_1 + I_2 + I_3 + I_4)$ and $(I_2 + I_3)/(I_1 + I_2 + I_3 + I_4)$ for each touch point. As shown in Fig. 2K, the contour lines near the edge of the panel were more curved than the lines near the center. Because one contour line was regarded to have the same β coordinate, the panel showed more distortion near the edge. The

distortion can be reduced by carefully modifying Eqs. 3 and 4 to nonlinear equations.

We performed experiments so as to prove the stability of the touch panel in a stretched state. We prepared a circular touch panel with a diameter of 4 cm and a homemade biaxial stretcher (24), as shown in Fig. 3A. We glued the panel to the biaxial stretcher and connected the panel to the controller board via Pt electrodes. As shown in Fig. 3B, the diameter of the touch panel was enlarged after stretching up to 12.5 cm, which corresponded to 1000% areal

strain. As shown in Fig. 3C, the touch panel can even be operated under highly stretched states. The related video clip is available in the supplementary materials (movie S2). The A1 currents of the touch panel at the undeformed and stretched states were measured with a touch point at the center in Fig. 3D. The baseline current increased from 6.68 to 9.16 μA , according to the stretch of the panel. The increase in the baseline current is suspected to be related to an expansion of the surface area of the gel panel (fig. S4), and the baseline current shows insensitivity to strain rate

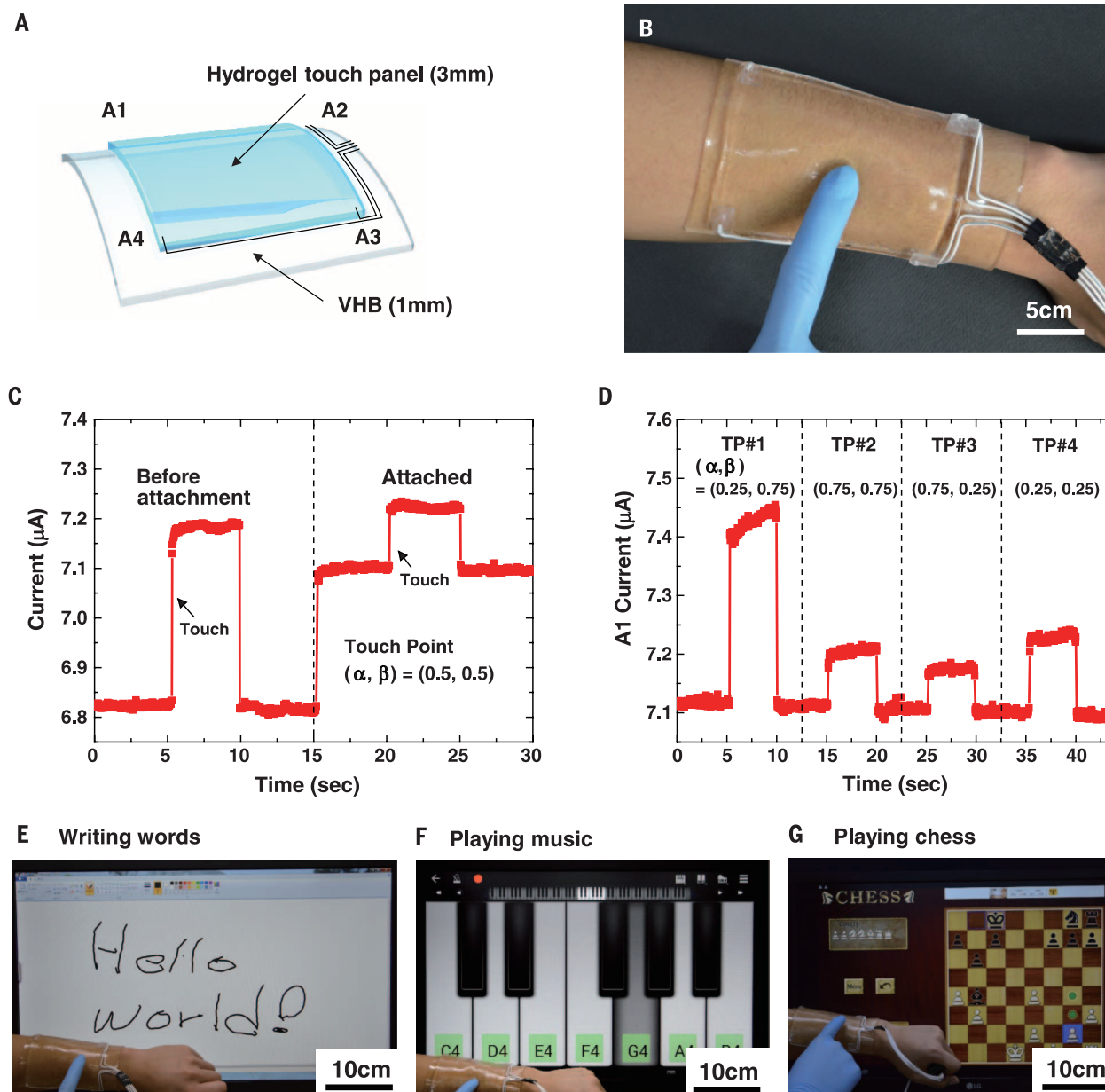


Fig. 4. An epidermal touch panel that is soft and transparent. (A) An epidermal touch panel was developed on a VHB substrate so as to insulate the panel from the skin and to mount the panel on a curved surface. (B) The touch panel was attached to an arm. (C) The A1 current of the panel was explored before and after attachment to an arm. (D) The position sensitivity after the attachment was tested by measuring the A1 currents at various touch points. (E to G) The epidermal touch panel is capable of detecting motions, such as tapping, holding, dragging, and swiping. Demonstrations such as writing words (E), playing music (F), and playing chess (G) are shown.

in the range of 0.2 to 1/min (fig. S5). When the stretched panel was touched, current was added to the baseline currents. The touching currents in a stretched state were 0.52 and 0.45 μA at $\lambda = 2$ and 3, respectively, which is similar to the current in the undeformed state (0.40 μA); the current response for a touch was not sacrificed by stretching. The current changes caused by touch during the stretching of a gel strip are shown in fig. S4. The baseline current gently increased, but the touching current was instantly increased when the strip was touched. Therefore, each current signal from stretching and touching can be distinguished

by its slope as well. Four touch points were sequentially investigated with a square touch panel at a stretch of $\lambda = 1.5$, as shown in Fig. 3E. The negative correlation between distance and current was maintained in a stretched state. The panel could be operated with an anisotropic deformation. As shown in fig. S7, the output figures were slightly shifted toward the undeformed edges after the deformation. However, when we applied an additional position calibration after the deformation, the shift was compensated.

Electromechanical stability was tested for the hydrogel through a cyclic loading test up to 100

cycles (Fig. 3, F to H). The measured resistance at various cycles is presented in Fig. 3H: 1, 10, 50, and 100 cycles. The resistance slightly increased as the cycles increased, showing a maximum variation of 25% when $\lambda = 3$. We suspect that the increase in resistance during a cyclic test may originate from water evaporation in the gel. The weight of the gel decreased from 9.6 to 7.54 g after the test. A uniaxial tensile test was performed with a hydrogel specimen, as shown in Fig. 3I. The hydrogel showed an elastic modulus of 12 kPa and an elongation of $\lambda = 11$.

A schematic design of an epidermal touch panel is shown in Fig. 4A. The epidermal touch panel was built on a 1-mm-thick VHB film (3M, Maplewood, MN) so as to insulate the panel from the body. Because VHB film was originally developed as an adhesive, the panel could be attached to an arm without using extra glues (Fig. 4B). The epidermal touch panel was fully transparent so that it could convey visual content behind the touch panel. Moreover, the panel was mechanically soft and stretchable so that a user is comfortable with movement while wearing it. The currents measured before and after attachment are plotted in Fig. 4C. The baseline currents increased after the attachment owing to a leakage of charges through the VHB substrate. The thicker insulating layer generated a smaller baseline current. The effect of thickness of the insulating layers on the baseline currents is shown in fig. S8. The sensitivity to touch decreased after the attachment; however, the touching current was still sufficient to be detected. As shown in Fig. 4D, we subsequently touched from TP#1 to TP#4 on the epidermal touch panel, and the current was measured with the A1 current meter. The correlation between the measured currents and the touched position was not influenced by the attachment. The epidermal touch panel could successfully perceive various motions, such as tapping, holding, dragging, and swiping. Thus, various applications can be easily managed by integrating the panel. As shown in Fig. 4, E to G, writing words (Fig. 4E), playing music (Fig. 4F), and playing chess (Fig. 4G) were accomplished via adequate motions on the epidermal touch panel (movies S3 to S6).

We have demonstrated a highly stretchable and transparent ionic touch panel. We used a PAAm hydrogel containing 2 M LiCl salts as an ionic conductor. We investigated the mechanism of position-sensing in an ionic touch panel with a 1D strip. The ionic touch strip showed precise and fast touch-sensing, even in a highly stretched state. We expanded the position-sensing mechanism to a 2D panel. We could draw a figure using the 2D ionic touch panel. The ionic touch panel could be operated under >1000% areal strain. An epidermal touch panel was developed based on the ionic touch panel. The epidermal touch panel could be applied onto arbitrarily curved human skin, and its use was demonstrated by writing words and playing the piano and games.

REFERENCES AND NOTES

- T. Young, U.S. patent 5,241,308 (1993).
- R. Aguilar, G. Meijer, *Proc. IEEE Sens.* **2**, 1360–1363 (2002).
- S. P. Hotelling, J. A. Strickon, B. Q. Huppi, U.S. patent 7,663,607 (2010).
- P. T. Krein, R. D. Meadows, *IEEE Trans. Ind. Appl.* **26**, 529–534 (1990).
- R. Adler, P. J. Desmares, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **34**, 195–201 (1987).
- M. R. Bhalla, A. V. Bhalla, *Int. J. Comput. Appl.* **6**, 12–18 (2010).
- D. Langley et al., *Nanotechnology* **24**, 452001 (2013).
- R. Bel Hadj Tahar, T. Ban, Y. Ohya, Y. Takahashi, *J. Appl. Phys.* **83**, 2631–2645 (1998).
- M. Vosgueritchian, D. J. Lipomi, Z. Bao, *Adv. Funct. Mater.* **22**, 421–428 (2012).
- Y. Xia, K. Sun, J. Ouyang, *Adv. Mater.* **24**, 2436–2440 (2012).
- L. Hu, W. Yuan, P. Brochu, G. Gruner, Q. Pei, *Appl. Phys. Lett.* **94**, 161108 (2009).
- Z. Wu et al., *Science* **305**, 1273–1276 (2004).
- J. Zang et al., *Nat. Mater.* **12**, 321–325 (2013).
- S. Bae et al., *Nat. Nanotechnol.* **5**, 574–578 (2010).
- L. Hu, H. S. Kim, J.-Y. Lee, P. Peumans, Y. Cui, *ACS Nano* **4**, 2955–2963 (2010).
- S. De et al., *ACS Nano* **3**, 1767–1774 (2009).
- C. F. Guo et al., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12332–12337 (2015).
- O. Akhavan, E. Ghaderi, *ACS Nano* **4**, 5731–5736 (2010).
- L. Ding et al., *Nano Lett.* **5**, 2448–2464 (2005).
- J.-Y. Sun et al., *Nature* **489**, 133–136 (2012).
- Y. Qiu, K. Park, *Adv. Drug Deliv. Rev.* **64**, 49–60 (2012).
- M. C. Darnell et al., *Biomaterials* **34**, 8042–8048 (2013).
- K. Obara et al., *Biomaterials* **24**, 3437–3444 (2003).
- C. Keplinger et al., *Science* **341**, 984–987 (2013).
- J. Y. Sun, C. Keplinger, G. M. Whitesides, Z. Suo, *Adv. Mater.* **26**, 7608–7614 (2014).
- C. Larson et al., *Science* **351**, 1071–1074 (2016).
- W. Pepper Jr., U.S. patent 4,293,734 (1981).
- H. Haga et al., *SID Symp. Dig. Tec.* **41**, 669–672 (2010).

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIP) (2015R1A5A1037668). J.-Y.S. and H.-H.L. acknowledge the support of the source technology and materials funded by the Ministry of Trade, Industry and Energy of Korea (MOTIE) (10052783).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/682/suppl/DC1
Materials and Methods
Figs. S1 to S11
References (29–31)
Movies S1 to S6

14 April 2016; accepted 19 July 2016
10.1126/science.aaf8810

SLEEP RESEARCH

Local modulation of human brain responses by circadian rhythmicity and sleep debt

Vincenzo Muto,^{1,2,3*} Mathieu Jaspard,^{1,2,3*} Christelle Meyer,^{1,2*} Caroline Kussé,^{1,2} Sarah L. Chellappa,^{1,2} Christian Degueldre,^{1,2} Evelyne Baletau,^{1,2} Anahita Shaffii-Le Bourdieu,^{1,2} André Luxen,^{1,2} Benita Middleton,⁴ Simon N. Archer,⁵ Christophe Phillips,^{1,2,6} Fabienne Collette,^{1,2,3} Gilles Vandewalle,^{1,2} Derk-Jan Dijk,^{5,†} Pierre Maquet^{1,2,7,†‡}

Human performance is modulated by circadian rhythmicity and homeostatic sleep pressure. Whether and how this interaction is represented at the regional brain level has not been established. We quantified changes in brain responses to a sustained-attention task during 13 functional magnetic resonance imaging sessions scheduled across the circadian cycle, during 42 hours of wakefulness and after recovery sleep, in 33 healthy participants. Cortical responses showed significant circadian rhythmicity, the phase of which varied across brain regions. Cortical responses also significantly decreased with accrued sleep debt. Subcortical areas exhibited primarily a circadian modulation that closely followed the melatonin profile. These findings expand our understanding of the mechanisms involved in maintaining cognition during the day and its deterioration during sleep deprivation and circadian misalignment.

Forgoing sleep and staying up at night, be it for professional or recreational reasons, is highly prevalent in modern societies (1). Acute sleep loss leads to deterioration of multiple aspects of cognition (2) and is associated with increased risk of human errors and health

hazards (3). These effects are often attributed to the mere lack of sleep. However, despite the progressive buildup of sleep pressure during wakefulness, human performance remains remarkably well preserved until wakefulness is extended into the biological night. This is attributed to a putative circadian alerting signal that increases during the day and reaches its peak in the early evening, close to the rise of melatonin concentration, to counter the mounting homeostatic sleep pressure (4–6). Cognition deteriorates rapidly and substantially when wakefulness is extended into the night and early morning hours. This is attributed to the accumulated sleep pressure and the dissipation of the circadian alerting signal (6, 7). Whether and how this interaction between homeostatic sleep pressure and circadian rhythmicity is represented at the regional brain level is not known. Single-time

¹GIGA-Cyclotron Research Centre–In Vivo Imaging, University of Liège, Liège, Belgium. ²Walloon Excellence in Life Sciences and Biotechnology (WELBIO), Liège, Belgium. ³Psychology and Cognitive Neuroscience Research Unit, University of Liège, Liège, Belgium. ⁴Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK. ⁵Sleep Research Centre, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK. ⁶Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium. ⁷Department of Neurology, CHU Liège, Liège, Belgium. *These authors contributed equally to this work. †These authors contributed equally to this work. ‡Corresponding author. Email: pmaquet@ulg.ac.be

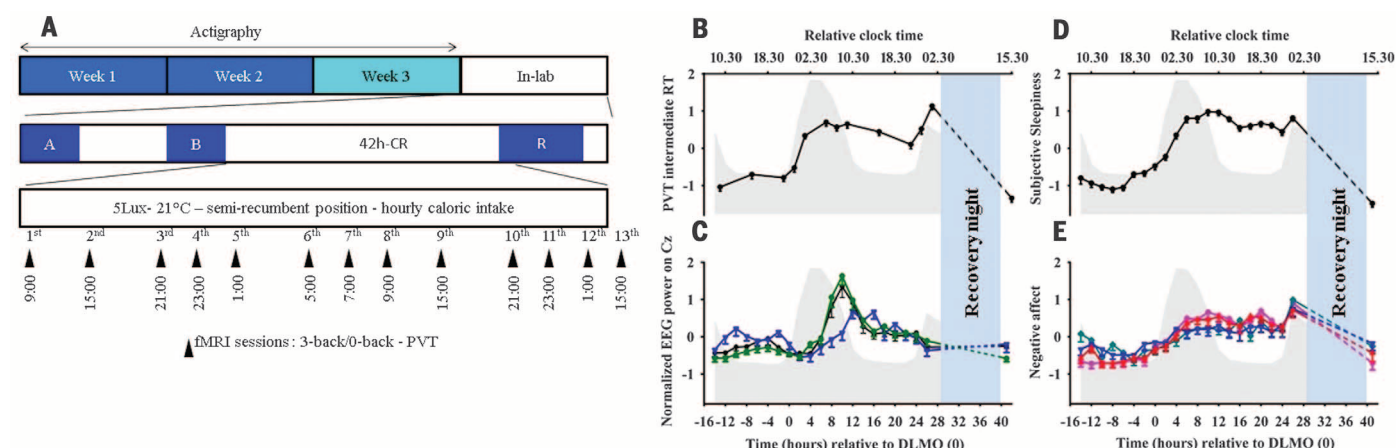


Fig. 1. Experimental protocol, behavioral, and physiological results. (A) Schematic representation of the experimental protocol. Actigraphy was recorded during 3 weeks prior to the laboratory study, which comprised an 8-hour adaptation night (A), an 8-hour baseline night (B), and a 12-hour recovery night (R). After baseline night, participants maintained wakefulness for 42 hours under constant conditions in dim light and in a semi-recumbent position (12). Note the clustering of the fMRI sessions in the morning (sessions 6 to 8) and evening (sessions 3 to 5 and 10 to 12). (B to E) Physiological and behavioral data, realigned to DLMO. The gray area illustrates the mean melatonin profile; the blue area represents the recovery sleep episode. All data are normalized z-scores, mean values \pm SEM. (B) PVT Intermediate reaction times varied

significantly across the 13 fMRI sessions ($F_{12, 366} = 55.52$, $P < 0.0001$). (C) Waking EEG power in delta (0.75 to 4.5 Hz, black line), theta (4.75 to 7.75 Hz, green line), and alpha (8 to 12 Hz, blue line) frequency bands. A main effect of time relative to DLMO was detected for delta ($F_{21, 577} = 8.44$, $P < 0.0001$), theta ($F_{21, 576} = 18.86$, $P < 0.0001$), and alpha power ($F_{21, 572} = 3.32$, $P < 0.0001$). (D) Subjective sleepiness varied significantly with time relative to DLMO ($F_{21, 629} = 58.51$, $P < 0.0001$). (E) Subjective status: stress (cyan; main effect of time relative to DLMO: $F_{21, 628} = 5.06$, $P < 0.0001$), anxiety (blue; $F_{21, 629} = 3.34$, $P < 0.0001$), happiness (red; $F_{21, 629} = 9.86$, $P < 0.0001$), and motivation (pink; $F_{21, 630} = 13.59$, $P < 0.0001$). Higher scores indicate higher levels of stress, anxiety, unhappiness, and demotivation.

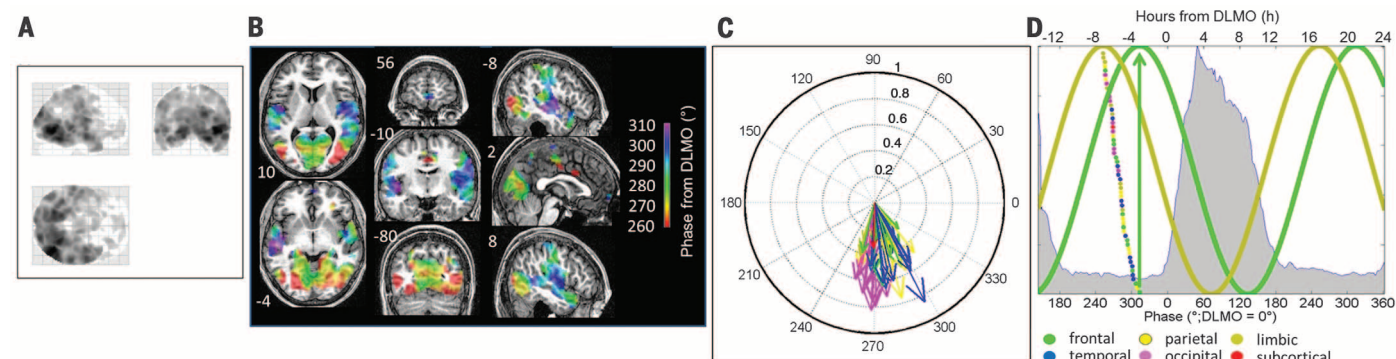


Fig. 2. PVT fMRI analysis 1. (A) Transparent brain display in Montreal Neurological Institute (MNI) space of areas showing significant responses with 24-hour periodicity estimated with the Sandwich Estimator method (21) ($P < 0.05$, FDR over the whole brain). (B) Circadian phase map of brain responses to PVT ($P < 0.05$, FDR over the whole brain). The local response phase is displayed according to the color scale ($^{\circ}$, DLMO = 0°) and overlaid over an individual normalized T1 MR scan. Coordinates are in millimeters along z, y, and x axes. (C) Polar representation of response phases ($^{\circ}$, DLMO = 0°). Arrow colors correspond to the color key in (D). (D) Predicted time courses of 24-hour period

responses expressed as phase and approximate hours from DLMO. Mean melatonin profile is shown in gray. Sine waves illustrate the earliest (beige, amygdala) and latest (green, inferior frontal gyrus) response timing. Between these two extreme phases, the staggered dots correspond to the timing of significant regional peak responses. These responses were grouped in six different areas according to the color code. Limbic phases ranged from 252° to 284° ; occipital, 255° to 270° ; frontal, 255° to 313° ; parietal, 255° to 308° ; temporal, 266° to 302° . Subcortical area consisted of left thalamus (coordinates, $-18 -30 -2$; phase, 268°).

point assessments of human brain responses to various cognitive tasks after acute sleep loss have demonstrated changes consistent with sleep loss' detrimental influence on brain information processing (8). However, an assessment of the time course of brain responses during sleep loss is currently not available.

We used repeated functional magnetic resonance imaging (fMRI) sessions to assess whether brain responses are modulated by circadian rhythmicity

during sustained wakefulness, and how circadian rhythmicity interacts with the sleep pressure accumulated with elapsed time awake and its dissipation during recovery sleep. Young healthy volunteers (17 men, 16 women; age 21.12 ± 1.7) stayed awake under constant environmental and behavioral conditions for a 42-hour period. The experiment started in the morning and covered two biological days, a full biological night, and the beginning of a second biological night. Brain

responses were assessed in 12 fMRI sessions clustered in the morning (hours ~ 05 to 09) and the evening/early night (hours 21 to 01), two periods characterized by rapid changes in the circadian modulation of cognitive performance. A 13th fMRI session took place after recovery sleep (Fig. 1A).

Circadian phase was determined from the central circadian pacemaker-driven melatonin rhythm (9), which showed a typical profile with low levels during the day and a sudden increase in the late

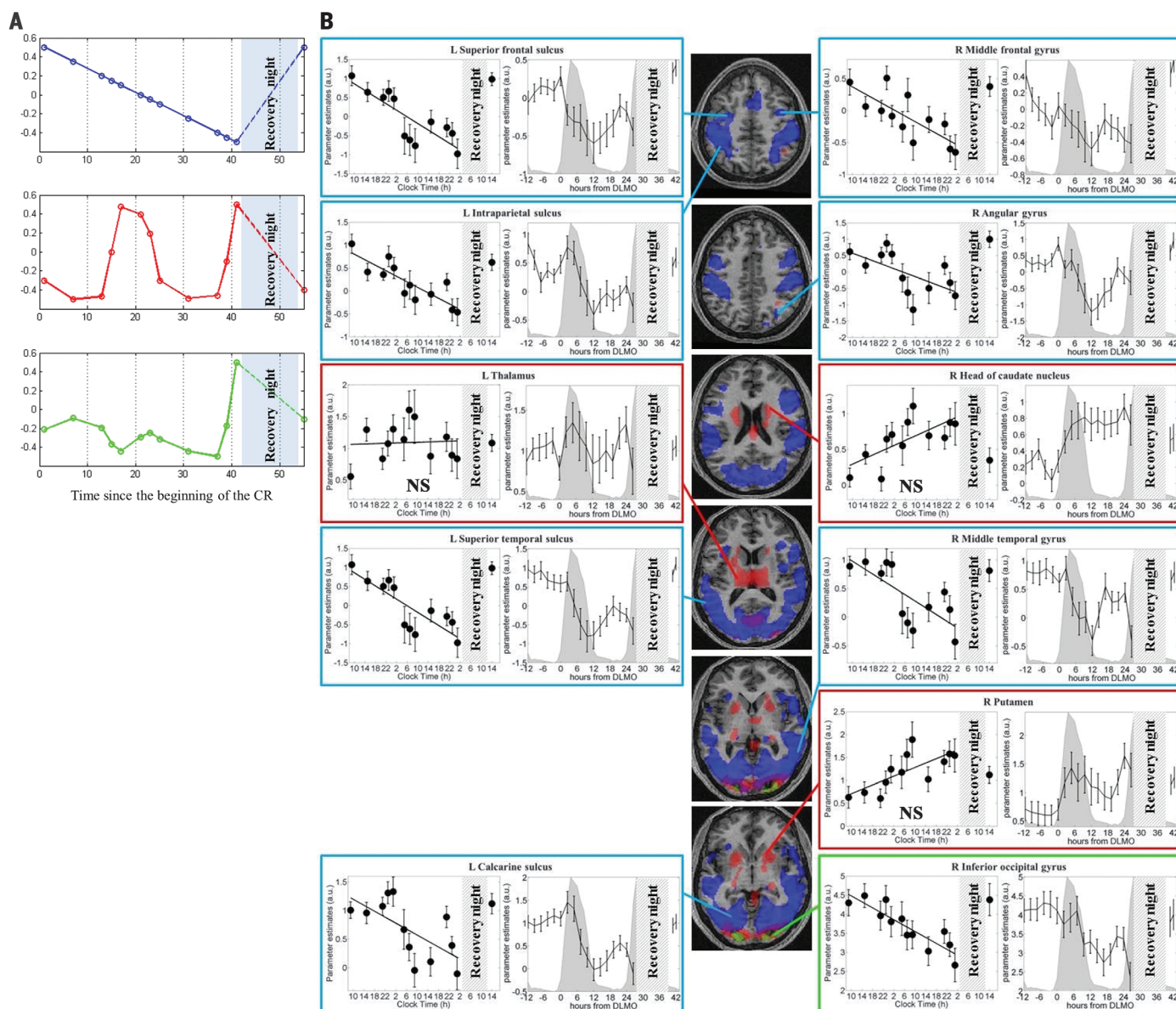


Fig. 3. PVT fMRI analysis 2. (A) Illustration of dimensionless fixed-effects fMRI contrasts testing (from top to bottom) a decrease in response with increasing sleep pressure during wakefulness and its recovery after sleep (blue), their fluctuation in association with mean melatonin level (red), and the interaction between these two factors (green). Note that the interaction is characterized by a steady level of response up to the evening sessions of day 3. (B) Images show significant effects of homeostatic sleep pressure (blue), circadian rhythmicity (red), and their interaction (green), displayed at $P < 0.05$ (FWE) over an individual normalized T1-weighted MR scan. Left and right

panels provide two different representations of the time course of brain responses, which were significant for sleep debt (blue border), circadian (red border), or the interaction (green border) contrasts. Irrespective of the contrast, beta estimates are plotted against clock time (left panels; linear regression is computed with respect to time awake during the sleep deprivation period) and time relative to DLMO (right panels; mean melatonin levels are shown in gray; activity estimates have been interpolated every 2 hours 24 min from hour -12 to hour $+28$). Coordinates are expressed in millimeters along z axis. NS, not significant.

evening hours [mean dim-light melatonin onset (DLMO), $22:33 \pm 00:09$ (SEM)]. Sleep during the 12-hour recovery night following the sleep deprivation was characterized by shorter sleep latency, increased sleep efficiency, total sleep time, and rapid eye movement (REM) and non-REM sleep, thereby confirming the increase in sleep pressure relative to baseline (tables S1 and S2).

Subjective sleepiness, negative affect, and delta and theta electroencephalographic (EEG) power increased with elapsed time awake and returned

to baseline after recovery sleep. These variables also showed a circadian modulation, with poorest ratings observed at the end of the biological night (at approximately 8:00 a.m.) and at the end of the sleep deprivation (at approximately 01:00 a.m.)—that is, after melatonin had risen again (Fig. 1, C to E).

During fMRI sessions, participants performed the psychomotor vigilance task (PVT) (10), which generated data on reaction times to pseudo-randomly occurring low-frequency stimuli (11). Reflecting

the effects of elapsed time awake and circadian phase, performance remained relatively stable during the first day, significantly declined after the first and second melatonin onsets, partially recovered during the second day, and returned to baseline after recovery sleep (Fig. 1B) (12).

A first fMRI voxelwise analysis identified any significant circadian periodicity in brain response profiles by combining two orthogonal regressors: 24-hour period sine and cosine waves adjusted to individual DLMO and computed for each individual

scan time. A significant circadian modulation was detected in a large set of cortical areas, involving nearly the whole cortical mantle [$P < 0.05$, false discovery rate (FDR) over the whole brain; Fig. 2A and table S3], with the exception of the dorsolateral prefrontal cortex.

Analyses of response phase showed significant variation across brain regions (Friedmann test, $\chi^2 = 30.13$, $df = 4$, $P = 4.60 \times 10^{-6}$) and a range spanning $\sim 250^\circ$ to $\sim 320^\circ$ relative to melatonin secretion onset (DLMO), with maximum responses occurring earlier in occipital and allocortical areas (amygdala and cingulate cortex) than in multimodal association areas (precuneus, temporal cortex, and prefrontal areas; Fig. 2, B to D). The predicted response profiles peaked during the subjective afternoon and reached their nadir in the second part of the night, up to the early morning hours, close to the offset of melatonin levels (Fig. 2D).

Although this analysis established a circadian modulation of regional brain responses, it assumed that the latter fluctuated as a sine wave—an assumption that does not correspond to actual time courses of most circadian biomarkers (12, 13). A second analysis evaluated the circadian modulation of PVT brain responses using an empirical marker of the circadian process: the mean melatonin level across volunteers (Fig. 3A, red). We simultaneously assessed whether brain responses to the PVT were modulated by accumulating sleep pressure (blue) and how circadian rhythm and homeostatic sleep pressure interact (green). Because no pure marker of homeostatic sleep pressure can be derived from empirical data obtained during sleep deprivation, its effect was modeled as monotonically decreasing with elapsed time awake and increasing after recovery sleep (12), as empirically observed (14) and usually modeled (15).

This second analysis indicated that the time course of responses significantly correlated with mean melatonin levels in a number of subcortical areas (midbrain, cerebellum, basal ganglia, and thalamus) and in a few cortical areas (primary sensorimotor cortices, occipital pole, and intraparietal sulcus), confirming a significant main effect of circadian rhythm in these regions (Fig. 3B, red areas, and table S5). In these areas, there was no significant effect of sleep debt. A significant negative effect of sleep debt was observed in a large set of cortical areas that spanned high-order association cortices of the frontal, parietal, insular, and cingulate cortices as well as visual and sensorimotor cortices (Fig. 3B, blue areas, and table S4). Their response pattern showed a decrease in response to elapsed time awake, with a return to baseline levels after recovery sleep (Fig. 3B, blue areas, and table S4). Their parameter estimates, adjusted to melatonin onset, also revealed a substantial (although not significant) circadian modulation, characterized by a rapid decrease in responses during the late subjective night or early subjective morning, around the melatonin offset. The circadian modulation appears more tightly in phase with melatonin levels in posterior areas than in more anterior areas, accounting for the significant interaction between sleep pressure and circadian rhythmicity observed in occipital poles

and thalamic areas (Fig. 3B, green areas, and table S6). Note that a transient increase of modeled cortical responses was no longer located in the “circadian” afternoon (as seen in analysis 1) but instead appeared immediately before the onset of melatonin secretion (Fig. 3B, response time courses relative to DLMO in frontal, parietal, and temporal cortices), a circadian time associated with low sleep propensity known as the wake maintenance zone (6).

Brain responses to an n -back task were also recorded during fMRI sessions. Executive responses (3-back > 0-back) in the bilateral anterior insula were significantly modulated by a circadian oscillation, synchronous to the melatonin rhythm [$P < 0.05$, familywise error (FWE) over the whole brain] (12) (fig. S1). This finding rules out a global task-independent circadian influence and suggests the influence of a local, region-specific, task-dependent circadian signal.

These findings reveal a pervasive effect of circadian rhythmicity and homeostatic sleep pressure on cortical responses during a sustained attention task. The interaction between circadian signals and sleep debt was formally proven in occipital areas, although inspection of response time courses suggests that both factors influence responses of many more cortical areas (Fig. 3B, right panels).

It appears that the respective influence of sleep debt and circadian rhythmicity is more balanced in posterior cortical areas, whereas sleep debt exerts a disproportionately larger influence in more anterior, associative areas. However, this generalization should be confirmed by further experimental data based on various cognitive tasks that are differentially affected by sleep loss and circadian rhythmicity (2).

More important, our results demonstrate a regional modulation of brain circadian rhythmicity. Several subcortical responses show a strong circadian modulation but no significant influence of sleep debt. By contrast, in most cortical areas, sleep pressure exerts a widespread negative influence on regional responses. This differential regulation of brain responses might explain the supposedly “compensatory” responses repeatedly reported in thalamic areas during sleep loss (8). In the morning after sleep deprivation (8 to 12 hours after DLMO, the typical assessment time point in fMRI studies of acute sleep deprivation), thalamic and striatal responses are indeed larger than in cortical areas. However, these strong thalamic responses might not reflect a compensation for the detrimental effects of accumulating sleep debt; they may merely indicate a dependency of cortical and subcortical response amplitude on the circadian phase (Fig. 3B). This observation highlights the importance of considering circadian phase when investigating the effects of sleep loss on brain mechanisms.

There is also a local modulation of cerebral circadian phase. This suggests that the circadian rhythmicity imposed by the master clock, located in the suprachiasmatic nucleus of the hypothalamus, can to some extent be locally altered, potentially in response to task-related requirements.

The mechanisms of this local modulation are unknown. Local changes in clock gene expression (16, 17) or posttranslational circadian mechanisms may be involved. Clock gene expression is sensitive to neuronal metabolic changes [e.g., redox state (18, 19)] and is altered in response to sleep debt (20).

These data demonstrate that sleep homeostasis and circadian rhythmicity affect brain responses, in accordance with current views on the regulation of sleep and waking performance. They also require a reformulation of these views to include the relative contributions of circadian rhythmicity and homeostatic sleep pressure to regionally specific (i.e., local) brain function. Our findings have implications for the understanding of the brain mechanisms underlying the maintenance of daytime cognitive performance and its deterioration, as observed in shift work, jet lag, sleep disorders, aging, and neurodegenerative diseases.

REFERENCES AND NOTES

1. E. Bixler, *Sleep Med.* **10** (suppl. 1), S3–S6 (2009).
2. J. C. Lo et al., *PLOS ONE* **7**, e45987 (2012).
3. S. M. Rajaratnam, J. Arendt, *Lancet* **358**, 999–1005 (2001).
4. D. J. Dijk, T. L. Shanahan, J. F. Duffy, J. M. Ronda, C. A. Czeisler, *J. Physiol.* **505**, 851–858 (1997).
5. J. K. Wyatt, A. Ritz-De Cecco, C. A. Czeisler, D. J. Dijk, *Am. J. Physiol.* **277**, R1152–R1163 (1999).
6. D. J. Dijk, C. A. Czeisler, *Neurosci. Lett.* **166**, 63–68 (1994).
7. D. J. Dijk, J. F. Duffy, C. A. Czeisler, *J. Sleep Res.* **1**, 112–117 (1992).
8. M. W. Chee, L. Y. Chuah, *Curr. Opin. Neurol.* **21**, 417–423 (2008).
9. P. Pevet, E. Challet, *J. Physiol. Paris* **105**, 170–182 (2011).
10. D. F. Dinges, J. W. Powell, *Behav. Res. Methods Instrum. Comput.* **17**, 652–655 (1985).
11. M. Basner, D. F. Dinges, *Sleep* **34**, 581–591 (2011).
12. See supplementary materials on Science Online.
13. C. Czeisler, O. M. Buxton, in *Principles and Practice of Sleep Medicine*, M. Kryger, T. Roth, W. Dement, Eds. (Elsevier, 2011), pp. 402–411.
14. J. K. Wyatt, C. Cajochen, A. Ritz-De Cecco, C. A. Czeisler, D.-J. J. Dijk, *Sleep* **27**, 374–381 (2004).
15. P. Achermann, *Aviat. Space Environ. Med.* **75** (suppl.), A37–A43 (2004).
16. X. Yu et al., *Curr. Biol.* **24**, 2838–2844 (2014).
17. L. E. Chun, E. R. Woodruff, S. Morton, L. R. Hinds, R. L. Spencer, *J. Biol. Rhythms* **30**, 417–436 (2015).
18. T. A. Wang et al., *Science* **337**, 839–842 (2012).
19. G. Asher et al., *Cell* **134**, 317–328 (2008).
20. P. Franken, D. J. Dijk, *Eur. J. Neurosci.* **29**, 1820–1829 (2009).
21. B. Guillaume, X. Hua, P. M. Thompson, L. Waldorp, T. E. Nichols, *Neuroimage* **94**, 287–302 (2014).

ACKNOWLEDGMENTS

Supported by Fonds National de la Recherche Scientifique (Belgium), Actions de Recherche Concertée of the Wallonia-Brussels Federation, University of Liège research funds, Fondation Médicale Reine Elisabeth, Fondation Simone et Pierre Clerdent, Bial Foundation, FEDER-Radiomed, and a Royal Society Wolfson Research Merit Award (D.-J.D.). We thank B. Guillaume for help in setting up the SWE analysis, E. Lambot and A. Golabek for assistance in data collection, and C. Schmidt for valuable feedback on the manuscript. Data are archived on CRC servers and available upon request. The authors report no conflict of interest.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/687/suppl/DC1
Materials and Methods
Supplementary Text
Tables S1 to S6
Figs. S1 and S2
References (22–51)

25 September 2015; accepted 20 June 2016
10.1126/science.aad2993

BRAIN MICROCIRCUITS

Imprinting and recalling cortical ensembles

Luis Carrillo-Reid,* Weijian Yang, Yuki Bando, Darcy S. Peterka, Rafael Yuste

Neuronal ensembles are coactive groups of neurons that may represent building blocks of cortical circuits. These ensembles could be formed by Hebbian plasticity, whereby synapses between coactive neurons are strengthened. Here we report that repetitive activation with two-photon optogenetics of neuronal populations from the visual cortex of awake mice builds neuronal ensembles that recur spontaneously after being imprinted and do not disrupt preexisting ones. Moreover, imprinted ensembles can be recalled by single-cell stimulation and remain coactive on consecutive days. Our results demonstrate the persistent reconfiguration of cortical circuits by two-photon optogenetics into neuronal ensembles that can perform pattern completion.

Cortical ensembles are groups of coactive neurons evoked by sensory stimuli (1–3) or motor behaviors (4–6) and likely constitute emergent building blocks of cortical function (7, 8). In the absence of external inputs, ongoing cortical ensembles resemble sensory evoked ones (9–11), as if the cortex has an imprinted representation of the world, implemented by groups of neurons with strong synaptic connectivity. Ensembles could result from Hebbian plasticity, whereby the connectivity between coactive neurons becomes strengthened (12). Optogenetic studies in which all expressing neurons and their axons are simultaneously photostimulated have dem-

onstrated Hebbian plasticity (13). However, the artificial generation of cortical ensembles with single-cell resolution has so far been experimentally difficult.

To do so, we used simultaneous two-photon calcium imaging and two-photon photostimulation (14, 15) in the primary visual cortex of head-fixed mice running on a treadmill. GCaMP6s signals of layer 2/3 neurons were imaged through a reinforced thinned-skull window, whereas CIV1-expressing neurons were optogenetically stimulated with a second two-photon laser (16) (Fig. 1, A and B).

Two-photon photostimulation of a neuronal population (fig. S1) evoked calcium transients reliably in a specific subset of neurons (Fig. 1, C to E). In vivo electrophysiological recordings demonstrated that population photostimulation evoked action potential bursts independently of the spatial

location of the neurons (fig. S1). Neurons responding to direct photostimulation were differentiated from other active neurons and from photostimulation light artifacts by their different temporal responses (fig. S1). This enabled us to distinguish photostimulated cells from those that became active because of the effects of photostimulation on the circuit.

Repeated optogenetic stimulation reliably recruited specific groups of neurons, generating an artificial “photoensemble” (i.e., a group of optically activated neurons). To quantify this, we used multidimensional population vectors to analyze population activity (17, 18) and found that photoensembles activate different populations of neurons than visual stimuli do (19), with only $20.17 \pm 9.4\%$ neurons in common (Fig. 2, A to D, and fig. S2). Although the number of ensembles was similar in both experimental conditions, photoensembles activated more neurons than visually evoked ones (Fig. 2E). Neurons belonging to photoensembles or visual ensembles had a widespread spatial distribution and were spatially intermingled (Fig. 2, F and G). Visual ensembles remained stable after population photostimulation (fig. S3), which indicates that repetitive photostimulation did not disrupt pre-existing cortical ensembles.

We noted that some photostimulated cells became coactive spontaneously (see below), as if the artificial photoensemble had been imprinted into the cortex. Moreover, the activation of a single cell was able to recall these imprinted ensembles (Fig. 3A), demonstrating pattern completion. Pattern completion (20) has been described in the hippocampus (21–23) and is a property of attractor neural networks (22, 24). Single-cell activation before population photostimulation did not produce considerable alterations of overall network activity (Fig. 3, B, C, and D, left, and fig.

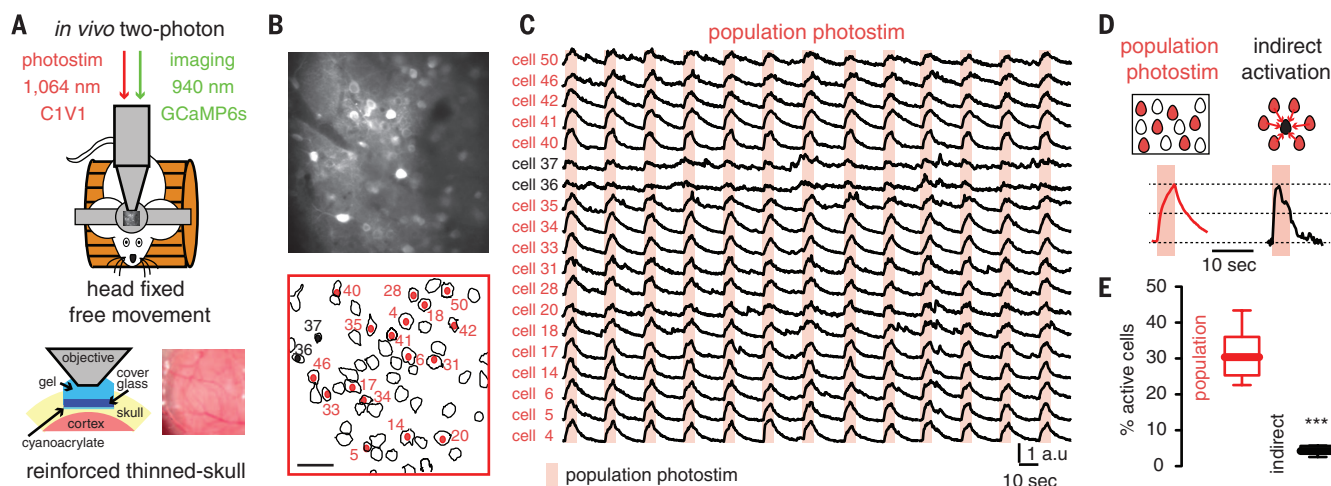


Fig. 1. Two-photon optogenetic photostimulation reliably activates specific neuronal populations. (A) Simultaneous two-photon imaging and two-photon optogenetic photostimulation were performed in layer 2/3 over the left primary visual cortex in awake head-fixed mice through a reinforced thinned-skull window. (B) Automatic contour detection of cortical neurons. Red cells denote neurons that reliably respond to optogenetic population photostimulation. Scale bar, 50 μ m. (C) Calcium transients of neurons ac-

tivated by population photostimulation (red) and neurons indirectly activated (black). a.u., arbitrary units. (D) Calcium transients from directly photostimulated neurons (left) differed from calcium transients evoked indirectly (right) by circuit activation. (E) Indirectly activated neurons represent a small percentage of the population ($n = 6$ mice; $***P = 0.0006$; Mann-Whitney test). Data in (E) are presented as box-and-whisker plots displaying median and interquartile ranges.

S4). However, after population photostimulation, photoactivation of selected members ($8 \pm 2.5\%$) of the imprinted ensemble (fig. S5) consistently

recalled an associated group of cells (Fig. 3, C and D, right). These recalled ensembles, evoked by single-cell stimulation, did not disrupt the

overall network activity and were interspersed in time with ongoing cortical ensembles (Fig. 3D, top). Though the number of ensembles during

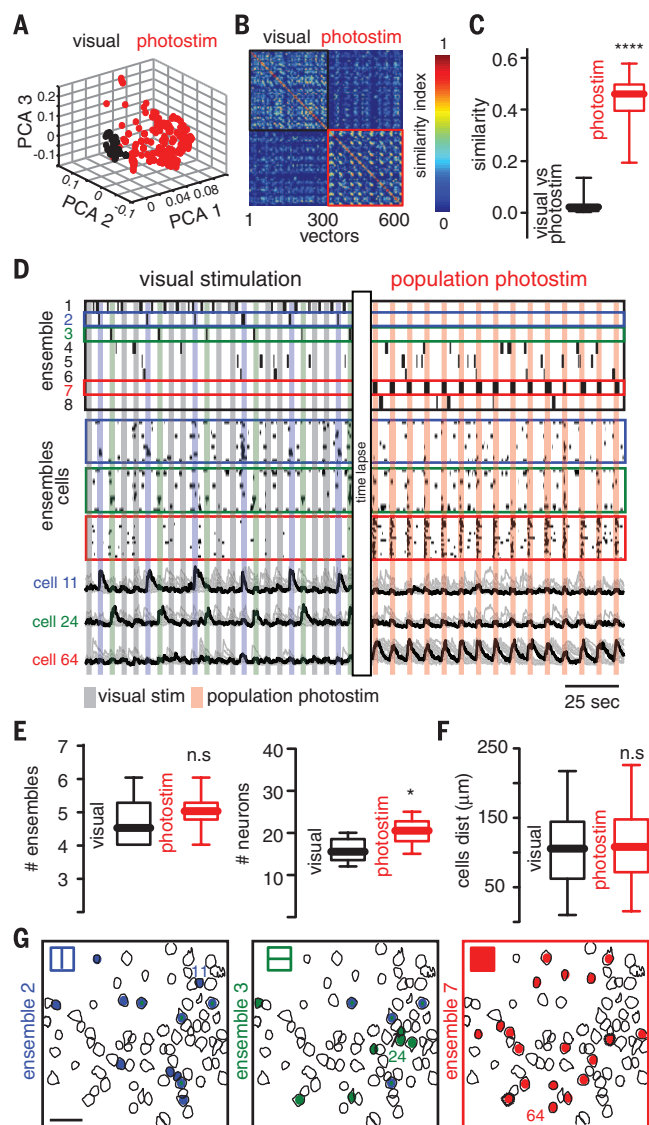


Fig. 2. Population photostimulation generates artificial cortical ensembles. (A) Principal component analysis (PCA) of population vectors evoked by visual stimuli (black) and optogenetic photostimulation (red). (B) Similarity map representing the angle between population vectors during visual stimuli (black) or population photostimuli (red). (C) Population similarity between visually and photostimulated evoked activity ($n = 6$ mice; $****P < 0.0001$). (D) Time-course of evoked cortical ensembles (top) aligned with raster plots representing the activity of visually evoked ensembles and photoensemble (middle) and calcium transients (bottom) of the most representative neurons of each ensemble. Colored boxes correspond to ensemble labels (blue, vertical stimulus; green, horizontal stimulus; red, photoensemble). (E) The total number of ensembles remained stable in both conditions [left; $n = 6$ mice; n.s. ($P = 0.4315$)]. The number of cells defining photoensembles is significantly higher than the number of neurons defining each visually evoked ensemble (right; $n = 6$ mice; $*P = 0.0446$). (F) Distance between all neurons belonging to each ensemble ($n = 6$ mice; n.s. ($P = 0.3720$)). (G) Spatial maps of cortical ensembles in both experimental conditions. Scale bar, $50 \mu\text{m}$. Data in (C), (E), and (F) are presented as box-and-whisker plots displaying median and interquartile ranges analyzed using the Mann-Whitney test.

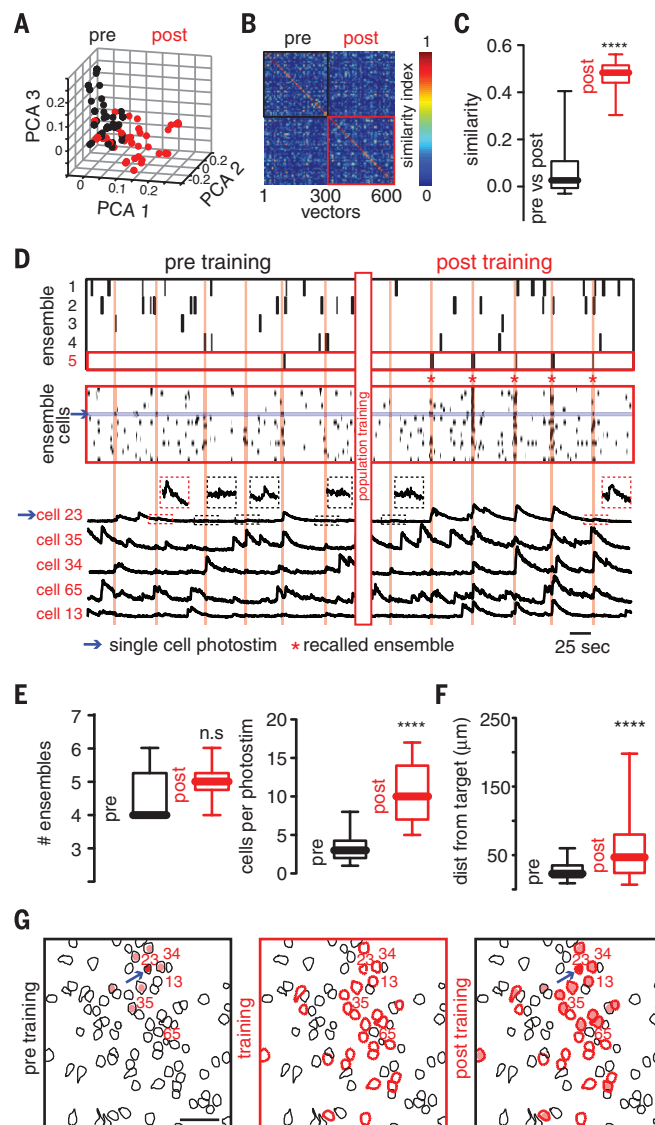


Fig. 3. Pattern completion of artificially imprinted ensembles. (A) PCA projection of population vectors during single-cell photostimulation before and after population training. (B) Similarity map of population vectors from ongoing cortical activity. (C) Single-cell photostimulation after population training enabled the recall of population vectors with high similarity ($n = 6$ mice; $****P < 0.0001$). (D) Time-course of cortical ensembles (top) aligned with a raster plot of all cells that belong to recalled ensembles (middle) and calcium transients (bottom) of representative neurons from recalled ensembles (red labels) before and after population training. Note how ensemble 5 (bottom, red) is reliably recalled by single-cell stimulation (cell 23) after training. (E) The number of ensembles before and after population training remains stable (left; $n = 6$ mice; n.s. ($P = 0.2259$)). After population training, single-cell photostimulation consistently recruits a group of neurons significantly larger than that existing under control conditions (right; $n = 6$ mice; $****P < 0.0001$). (F) After population training, the distance between the target cell and the activated neurons is increased ($n = 6$ mice; $****P < 0.0001$). (G) Spatial maps of neurons recruited by single-cell photostimulation before (left), during (middle), and after population training (right). The arrow indicates a stimulated neuron. Scale bar, $50 \mu\text{m}$. Data in (C), (E), and (F) are presented as box-and-whisker plots displaying median and interquartile ranges analyzed using the Mann-Whitney test.

single-cell photostimulation before and after population training remained stable (Fig. 3E), single-cell photostimulation after population training reliably enabled the recall of a specific group of neurons that was not coactive before; this occurred $64.5 \pm 12.63\%$ of the time (Fig. 3E). The spatial location of neurons in recalled ensembles had a broader distribution than the occasional neurons that were indirectly activated before population training (Fig. 3, F and G). Also, after population training, the number of calcium

transients in nonphotostimulated neurons remained constant whereas it increased in photostimulated neurons (Fig. 4, A and B), ruling out the possibility that population photostimulation changed the basal level of activity in the whole network. This modification of the functional connectivity between photostimulated neurons required a minimal number of trials (Fig. 4C), indicating that the observed changes were driven by an alteration in the circuit triggered by repeated photostimulation of a specific population of neurons.

To investigate whether imprinted ensembles were persistently integrated in ongoing cortical activity, we imaged the same area on consecutive days. Single-cell photostimulation still enabled the recall of previously imprinted ensembles on consecutive days (fig. S6). Our analysis of ongoing activity from nonphotostimulated (Fig. 4D, left) and photostimulated neurons showed that imprinted ensembles recurred spontaneously, even on consecutive days and after additional population photostimulation (Fig. 4D, right). Although

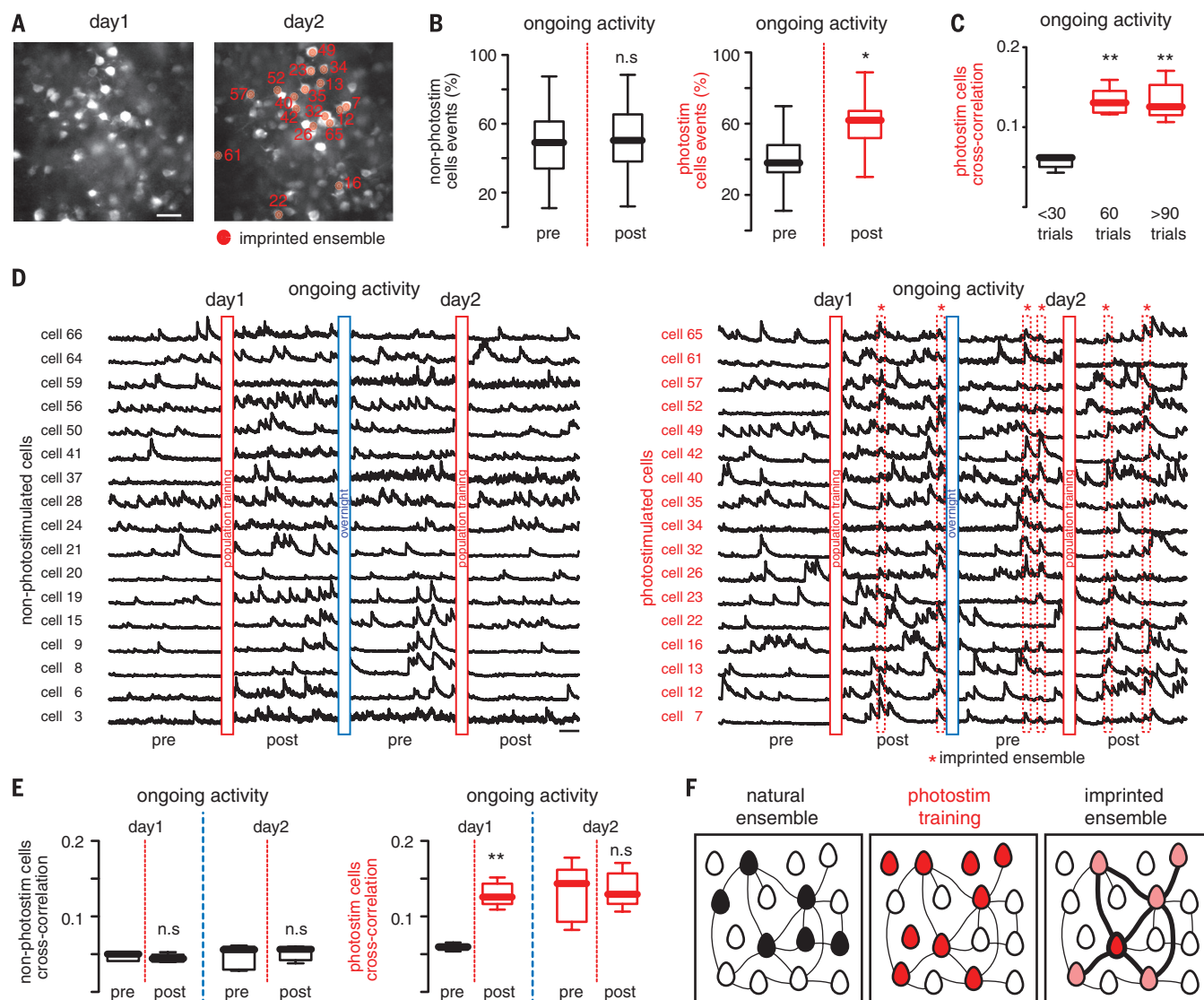


Fig. 4. Imprinted ensembles persist after consecutive days. (A) Images showing the same optical field on two different days. Scale bar, 50 μm . (B) The percentage of events during ongoing activity of nonphotostimulated cells remains stable (left; $n = 5$ mice; n.s. $P = 0.5664$; Wilcoxon matched-pairs signed rank test), whereas photostimulated cells increased their activity (right; $n = 5$ mice; $*P = 0.0147$; Wilcoxon matched-pairs signed rank test) after population training. The red dashed lines denote the occurrence of population training. (C) Enhancement of cross-correlation between photostimulated cells depends on the number of training trials ($n = 5$ mice; $**P = 0.0092$; Kruskal-Wallis test). (D) Calcium transients of nonphotostimulated neurons (left) and photostimulated neurons (right) during ongoing cortical activity on two different days. Imprinted ensembles recur spontaneously

on consecutive days (indicated by the dotted red boxes), even after a second photostimulation training. (E) Cross-correlation between nonphotostimulated neurons (left; $n = 5$ mice; day 1: n.s. $P = 0.5476$; day 2: n.s. $P = 0.8413$; Mann-Whitney test) and photostimulated neurons (right; $n = 5$ mice; day 1: $**P = 0.079$; day 2: n.s. $P = 1$; Mann-Whitney test) during ongoing activity on consecutive days. (F) Proposed model: Population photostimulation enhances the functional connectivity between responsive neurons. Line widths represent the strength of the functional connectivity between neurons. Black, neurons belonging to preexisting ensembles; red, photostimulated neurons; pink, recalled neurons during pattern completion. Data in (B), (C), and (E) are presented as box-and-whisker plots displaying median and inter-quartile ranges.

cross-correlations between nonresponsive neurons were not altered (Fig. 4E, left), they were increased between photostimulated neurons and remained stable the next day (Fig. 4E, right). Thus, optogenetic activation of identified neurons enhanced their local functional connections for at least 1 day (Fig. 4F).

Recalled ensembles shared similar characteristics—such as number of neurons and spatial distribution—with ongoing ensembles (fig. S7), but the mean distance between active neurons was shorter (fig. S7D), which indicates that the effect of the photostimulation is local. Recalled ensembles often had neurons that did not belong to ongoing ensembles (fig. S7, D and E), demonstrating that recalled ensembles are indeed novel and not just dormant preexisting ensembles. However, given that cortical connections are likely not in a tabula rasa state, we expect that imprinted ensembles may recruit segments of physiologically relevant circuit motifs (Fig. 4F).

Previously, electrical or optogenetic stimulation (25) has been used to show that coactivation of neuronal groups can produce physiologically relevant behaviors (13, 26). Here, we show the possibility of training individual neurons to build artificial neuronal ensembles (13), which then become spontaneously active (Fig. 4D, right). Our results are consistent with the finding that neurons responding to similar visual stimuli have a higher interconnectivity (27), as well as with the similarity between visually evoked and spontaneous ensembles (9). In both cases, recurrent coactivation of a neuronal group would enhance functional connectivity, imprinting ensembles into the circuit.

More than 60 years ago, Hebb proposed that repeated coactivation of a group of neurons might create a memory trace through enhancement of synaptic connections (12). Because of technical limitations, this hypothesis has been difficult to test with single-cell resolution in awake animals. By combining novel imaging and photostimulation techniques (14, 15) and analytical tools (19), our work can be interpreted as a confirmation of the Hebbian postulate and as a demonstration that cortical microcircuits can perform pattern completion.

REFERENCES AND NOTES

- B. M. Kampa, M. M. Roth, W. Göbel, F. Helmchen, *Front. Neural Circuits* **5**, 18 (2011).
- K. Ohki, S. Chung, Y. H. Ch'ng, P. Kara, R. C. Reid, *Nature* **433**, 597–603 (2005).
- J. Sawinski et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19557–19562 (2009).
- M. M. Churchland et al., *Nature* **487**, 51–56 (2012).
- A. J. Peters, S. X. Chen, T. Komiyama, *Nature* **510**, 263–267 (2014).
- V. Y. Cao et al., *Neuron* **86**, 1385–1392 (2015).
- V. B. Mountcastle, *Brain* **120**, 701–722 (1997).
- R. Yuste, *Nat. Rev. Neurosci.* **16**, 487–497 (2015).
- J. E. Miller, I. Ayzenshtat, L. Carrillo-Reid, R. Yuste, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4053–E4061 (2014).
- T. Kenet, D. Bibitchkov, M. Tsodyks, A. Grinvald, A. Arieli, *Nature* **425**, 954–956 (2003).
- A. Luczak, P. Barthó, K. D. Harris, *Neuron* **62**, 413–425 (2009).
- D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory* (Wiley, 1949).
- J. P. Johansen et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12692–12697 (2010).
- A. M. Packer, L. E. Russell, H. W. Dalgleish, M. Häusser, *Nat. Methods* **12**, 140–146 (2015).
- J. P. Rickgauer, K. Deisseroth, D. W. Tank, *Nat. Neurosci.* **17**, 1816–1824 (2014).
- P. J. Drew et al., *Nat. Methods* **7**, 981–984 (2010).
- S. L. Brown, J. Joseph, M. Stopfer, *Nat. Neurosci.* **8**, 1568–1576 (2005).
- L. Carrillo-Reid et al., *J. Neurophysiol.* **99**, 1435–1450 (2008).
- L. Carrillo-Reid, J. E. Miller, J. P. Hamm, J. Jackson, R. Yuste, *J. Neurosci.* **35**, 8813–8828 (2015).
- M. R. Hunsaker, R. P. Kesner, *Neurosci. Biobehav. Rev.* **37**, 36–58 (2013).
- R. C. O'Reilly, J. L. McClelland, *Hippocampus* **4**, 661–682 (1994).
- E. T. Rolls, A. Treves, *Prog. Brain Res.* **102**, 335–341 (1994).
- E. T. Rolls, R. P. Kesner, *Prog. Neurobiol.* **79**, 1–48 (2006).
- J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554–2558 (1982).
- A. Jackson, J. Mavoori, E. E. Fetz, *Nature* **444**, 56–60 (2006).
- X. Liu et al., *Nature* **484**, 381–385 (2012).
- H. Ko et al., *Nature* **473**, 87–91 (2011).

ACKNOWLEDGMENTS

We thank our laboratory members for help and virus injections, A. Fairhall for comments, and the Stanford Neuroscience Gene Vector and Virus Core for AAVdj virus. This work was supported by

the National Eye Institute (grants DP1EY024503 and R01EY011787), National Institute of Mental Health (grants R01MH101218, R01MH100561, R41MH100895, and R44MH109187), and Defense Advanced Research Projects Agency (grant SIMPLEX N66001-15-C-4032). Y.B. holds a fellowship from Uehara Memorial Foundation. W.Y. holds a Career Award at the Scientific Interface from Burroughs Wellcome Fund. This material is based on work supported by, or in part by, the U.S. Army Research Laboratory and the U.S. Army Research Office (contract W911NF-12-1-0594, Multidisciplinary University Research Initiative). We declare no competing financial interests. Author contributions: L.C.-R. and R.Y. came up with the concept for this work. L.C.-R., D.S.P., W.Y., and R.Y. designed the methodology. L.C.-R., W.Y., and Y.B. carried out the investigation. L.C.-R. wrote the original draft. L.C.-R. and R.Y. reviewed and edited the paper. D.S.P., W.Y., and L.C.-R. provided the resources for this work. R.Y. acquired funding. All of the data are archived in the NeuroTechnology Center at Columbia University.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/691/suppl/DC1
Materials and Methods
Figs. S1 to S7
References (28–40)

24 March 2016; accepted 20 July 2016
10.1126/science.aaf7560

ECONOMIC POLICY

The impact of homelessness prevention programs on homelessness

William N. Evans,^{1,2,3} James X. Sullivan,^{1,3*} Melanie Wallskog⁴

Despite the prevalence of temporary financial assistance programs for those facing imminent homelessness, there is little evidence of their impact. Using data from Chicago from 2010 to 2012 ($n = 4448$), we demonstrate that the volatile nature of funding availability leads to good-as-random variation in the allocation of resources to individuals seeking assistance. To estimate impacts, we compare families that call when funds are available with those who call when they are not. We find that those calling when funding is available are 76% less likely to enter a homeless shelter. The per-person cost of averting homelessness through financial assistance is estimated as \$10,300 and would be much less with better targeting of benefits to lower-income callers. The estimated benefits, not including many health benefits, exceed \$20,000.

Over 2 million people experience homelessness each year in the United States (1). Historically, the primary approach to combating homelessness has been to provide emergency shelters or transitional housing services to those who are already homeless. More recently, policy-makers have increased their focus on homelessness prevention efforts. One of the most common prevention strategies is to provide temporary financial assistance to

people facing eviction in order to keep them in their residences. In the United States, 93% of households live in an area that has such a program, and these programs receive over 15 million calls a year (2). Despite the prevalence of these efforts, there is little evidence about the extent to which they actually prevent homelessness (3, 4).

Here we examine the effectiveness of temporary financial assistance by using data from the Homelessness Prevention Call Center (HPCC) in Chicago, which processes about 75,000 calls annually. Chicago residents at risk of becoming homeless can call 311 to request temporary financial assistance for rent, security deposits, or utility bills. These callers are routed to the HPCC, which is a centralized processing center that screens callers for eligibility and connects eligible callers with local funding agencies.

¹Department of Economics, University of Notre Dame, Notre Dame, IN 46556, USA. ²National Bureau of Economic Research, Cambridge, MA 02138, USA. ³Wilson Sheehan Lab for Economic Opportunities, Notre Dame, IN 46556, USA.

⁴Department of Economics, Stanford University, Stanford, CA 94305, USA.

*Corresponding author. Email: jsullivan4@nd.edu

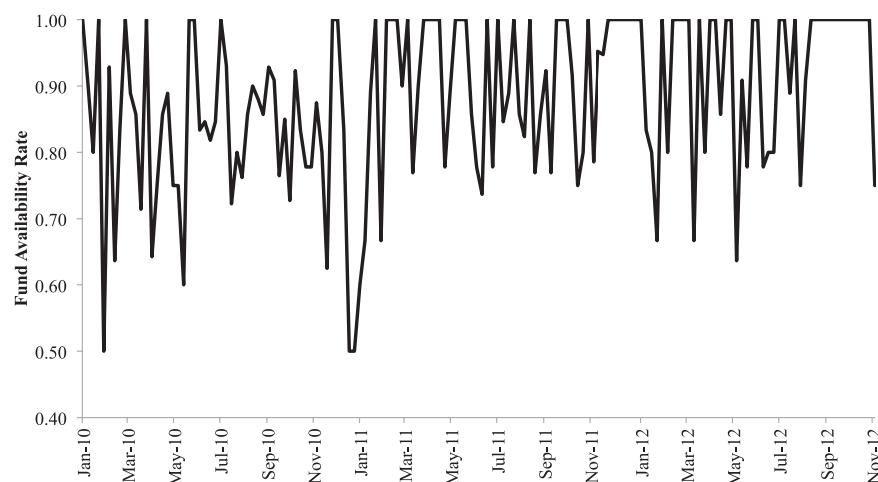


Fig. 1. Fund availability rate by week. Shown is the fraction of eligible callers in our homogenous sample that were referred to financial assistance each week from 2010 to 2012. The homogeneous sample includes all eligible callers who were seeking rent assistance with need amounts between \$301 and \$900, were nonveterans, were not receiving housing subsidies, were not requesting more than 1 month's rent, had a social security number, had a family income below twice the poverty line, and were not homeowners ($n = 1431$).

At the beginning of each call to the HPCC, information and referral specialists collect detailed information to determine whether the client is eligible for financial assistance. The specialist refers those who are not eligible for financial assistance to nonfinancial support services. General eligibility is based on four criteria: (i) The client must be able to demonstrate self-sufficiency after they receive assistance; (ii) the client must have an eligible crisis (e.g., job loss or medical emergency) that has led to the need for assistance; (iii) the client must face imminent risk of homelessness or utility shutoff; and (iv) the current crisis must be solvable by the financial assistance.

For eligible callers, the specialist then looks to see whether there is a delegate agency that is offering funds at that time. These agencies have additional fund-specific restrictions beyond those imposed by the general eligibility rules, such as the request type (rent, mortgage, security deposits, or utility bills), the need amount, veteran status (a few funds are restricted to veterans), receipt of housing subsidies (some funds will not assist those who receive Section 8 vouchers), the number of months of unpaid rent (some funds will pay for only 1 month), whether the caller's income is more than twice the poverty line, whether the caller has a valid social security number, and whether the caller is a homeowner.

Not all eligible callers are referred to funds. Funding for financial assistance varies unpredictably over time. The availability of funding on any given day depends on many factors. For example, some delegate agencies require that callers meet with a financial counselor before funds are dispersed, and the caller will not be referred if an interview slot is not available at

the time of the call. For some agencies, there are only a fixed number of appointments available each week or month, but new interview slots might become available through cancellations. Variation in funding can also occur as a result of inconsistent and unpredictable funding streams from local or state programs.

The variation in the availability of funding is evident in Fig. 1, which shows the fund availability rate (the fraction of eligible callers that are referred for funding) at the HPCC by week from 2010 through 2012. To ensure that this variation is not due to changes over time in caller characteristics, we focused on a subset of callers, which we call the "homogeneous" subsample, whose characteristics affecting access to specific funds were the same. For example, all callers in this subsample were requesting rent assistance, and all had need amounts between \$301 and \$900. As Fig. 1 shows, there was considerable variation in the likelihood that such callers were referred for assistance. For some weeks, all eligible callers with these characteristics were referred to funds. But for most weeks, only a subset of these callers was referred, and for two of these weeks only half were referred. We used this variation in fund availability to determine the impact of financial assistance on homelessness. One concern could be that this variation may be correlated with caller characteristics that directly influence homelessness because certain types of callers may have information about fund availability. However, from the perspective of the eligible client, the availability of funds is difficult to predict (supplementary materials, section 1.3, and tables S1 and S2).

The empirical analysis for this study relies on administrative data about callers seeking

temporary financial assistance that are routed to the HPCC (supplementary materials, materials and methods). The main sample used for this study was drawn from all calls to the HPCC from 20 January 2010 to 4 December 2012. We restricted our sample to requests for rent or security deposits because those who call the HPCC seeking assistance with utility bills are highly unlikely to enter a homeless shelter (supplementary materials, section 1.2), which is our primary outcome of interest. We also restricted the sample to the first call to the HPCC from an individual. It is common for some callers to contact the HPCC multiple times. Our concern was that subsequent calls may not be exogenous: The characteristics associated with these calls may be correlated with both the availability of funds and the likelihood of entering a homeless shelter. Our main sample for analysis thus included 4448 calls—3574 for rent and 874 for security deposits—and 58.2% of these callers were referred to funds (table S3).

If the availability of funds were purely random, one could determine the impact of financial assistance by comparing outcomes for eligible individuals who call the HPCC when funds are available with outcomes for those who call when funds are not available. In practice, however, not all eligible callers have the same likelihood of being referred to funds because of the fund-specific restrictions mentioned above. For example, callers who requested between \$301 and \$900 of assistance were more likely to call when funds were available because more funds cater to these need amounts (table S4). Therefore, for our balancing tests of baseline characteristics, we controlled for the factors that are known to affect fund availability by estimating regressions of individual characteristics on a dummy variable for fund availability and these factors. To account for patterns in call volume, we also included seasonal controls (supplementary materials, section 1.4).

The results from our balancing tests are reported in Table 1 for both the homogeneous subsample and our main sample of eligible callers seeking assistance with rent or security deposits. In the first and third columns, we present the means for observable characteristics for our control group—callers who are not referred for funding. In the second and fourth columns, we report the difference in means between the treatment and control groups, as measured by the coefficient on the availability-of-funds indicator. For the homogeneous subsample, no additional controls were included in the regression, so this estimate is the actual raw difference in means. For the main sample, this estimate is from a regression that included the other controls mentioned above. The results indicate that those who call for rent or security deposit assistance when funding is not available are very similar to those who call when funding is available. In nearly all cases, we failed to reject the null hypothesis that the characteristics are the same at $P < 0.05$ (5). For the characteristics where we did reject the null hypothesis,

the differences in means were small. As an example, for the main sample, those calling when funds were available made \$52 less in monthly income than those calling when funds were not available, a 4% difference in income. The statistically significant differences indicate that those calling when funds are available are more likely to enter a shelter, independent of the availability of funds, which would bias against our main finding that funding reduces homelessness. Additional evidence of the balance in characteristics is available in table S5.

To assess homelessness outcomes, we linked the HPCC data to administrative data from the Homeless Management Information System (HMIS) about entries into and exits from housing facilities for the homeless in Chicago. These data cover roughly two-thirds of homeless beds in Chicago (supplementary materials, section 1.1, and table S6). In the results presented below, we focus on these key measures of homelessness: whether a caller enters a homeless shelter within 3 or 6 months after the call and the number of days spent in a shelter during the first 6 months after the call.

We estimated the intention-to-treat (ITT) effect of being referred to financial assistance on the likelihood that an individual enters a homeless shelter (supplementary materials, materials and methods). These ITT estimates capture the difference in homelessness outcomes between those who call when funds are available and those who call when funds are unavailable. They measure the impact of fund availability, not actual receipt of financial assistance (6, 7).

The results of our analysis of the impact of fund availability on homelessness are presented in Table 2 for our main sample, the homogeneous subsample, and other subsamples (estimates of the coefficients on other covariates are given in table S8, and alternative specifications that demonstrate robustness are given in tables S7 and S9). For our main sample, fund availability led to a 1.4-percentage-point decrease in the probability of entering a shelter within 3 months of the call and a 1.6-percentage point decrease in the probability of entering a shelter within 6 months (8). Both of these estimates are statistically significant at $P < 0.05$. These effects represent an 88% decline in the likelihood of becoming homeless after 3 months and a 76% decline in the likelihood after 6 months. The results for days spent in a shelter are consistent with those for whether a caller enters a shelter. Calling when funds were available reduced the time spent in a shelter over the next 6 months by 2.6 days (or 84%), and nearly all of the effect on days spent in a shelter can be accounted for by whether a caller entered a shelter rather than by the length of the stay. This suggests that fund availability has only a small effect on the length of time one spends in a shelter, conditional on entry. That the impact of fund availability falls predominantly on whether or not one enters a shelter is not surprising, given that the intervention is designed to keep people in their own residences.

The impact of calling when funds are available is comparable for the homogeneous subsample (Table 2). The magnitudes of the effect are larger (in absolute value), but in percentage terms, the impact is slightly smaller, and these estimates are not statistically significant at $P < 0.05$. We found a similar impact for the subsamples of callers seeking rent or security deposit assistance; in nearly all of these cases, the effect of fund availability is statistically significant at $P < 0.05$ (Table 2). For example, for those seeking rent assistance, fund availability led to a 1.5-percentage-point decrease in the probability of entering a shelter within 6 months, a decline of 71%.

To examine how the impact changes over time since the call, we re-estimated the ITT effect with the dependent variable in the model being whether the caller checked into a shelter within τ months after the call, where τ ranges from 1 to 12. Figure 2 reports the main point estimates from these specifications along with the 95% confidence interval for our main sample. These results show that calling when funds are available has an immediate impact on homelessness, which is not surprising given that most eligible callers face imminent risk of homelessness: Most have already been served an eviction notice. The effect within 2 months is slightly larger than the effect within 1 month, but the magnitude of the effect changes little at longer follow-up periods. The estimated impact within 12 months (a decline of 1.7 percentage points) is similar to the estimated impact within 2 months (a decline of 1.5 percentage points). In all cases, the estimates are statistically significant, and the results are similar for the subsamples of callers

seeking assistance with rent (fig. S1) or security deposits (fig. S2).

Analyses for other subgroups reveal considerable heterogeneity in the effects of financial assistance. As shown in Table 2, the effects differ noticeably by income. For those seeking rent assistance whose incomes were below or equal to the median income for the sample (below ~90% of the federal poverty line), calling when funds were available reduced the likelihood of entering a shelter within 6 months by 2.2 percentage points (88%). There is little evidence of an effect of financial assistance on homelessness for those with above-median income among eligible callers. Other subgroups that appear more likely to benefit from financial assistance include individuals (single adults), males, callers younger than 30, and those calling in the winter (table S10) (9).

A common criticism of programs that aim to prevent homelessness through financial assistance is that emergency funds tend to be poorly targeted; resources go to those who would not end up homeless even in absence of the assistance (4, 10, 11). Thus, the temporary financial assistance may crowd out other resources for avoiding homelessness. The evidence from our sample of callers to the HPCC is consistent with this argument. Many HPCC callers who were not referred to financial assistance found a way to avoid homelessness even though they were facing eviction from their residence at the time of the call (12). Despite this evidence of considerable crowd-out, given the high cost of homelessness to individuals and society, even a small overall reduction in homelessness may be cost-beneficial. Thus, it is important to know whether the value to individuals and society of the unmeasured

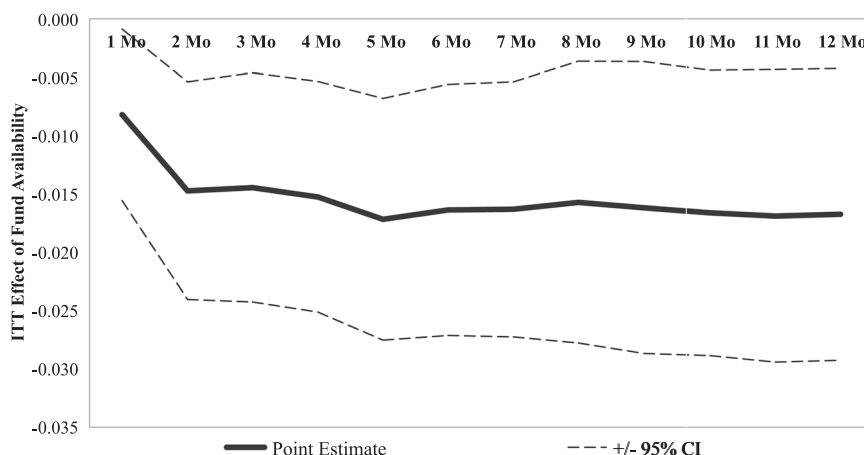


Fig. 2. ITT effects of fund availability on shelter admittance 1 to 12 months after the call. We estimated 12 separate regressions of shelter entry 1 to 12 months (mo) after the call on an availability-of-funds indicator, demographic characteristics of the caller, ZIP code-level characteristics, fund-specific restriction variables, and seasonal control variables. The solid line plots the coefficients on the availability-of-funds indicator in these regressions (e.g., -0.015 denotes a 1.5-percentage-point decline in shelter entry); estimates of the coefficients on other covariates are given in table S8. The dashed lines denote the 95% confidence interval (CI) calculated using standard errors clustered at the ZIP code level. The estimates for the effect 1 to 6 months after the call are based on the main sample ($n = 4448$). Beyond 6 months, we lose some information because we do not have data about shelter entry beyond 6 months for all callers. Thus, the sample size decreases with each month beyond 6 months, with the 12-month shelter entry estimates based on a sample of 3892 callers.

benefits of homelessness reduction is large enough to justify the documented costs.

The cost of reducing homelessness through rent assistance includes the operating costs of the call center and delegate agencies, as well as the cost of the financial assistance. Altogether, we estimate these costs to be about \$720 per caller referred. Given our main estimates of the effect of calling the HPCC when funding is available, we calculate the cost per homeless spell averted to be about \$10,300 (supplementary materials, section 2.4). This cost per home-

less spell averted would be lower if the program were better targeted toward those for whom financial assistance is more effective at reducing homelessness. Our results indicate that the per-person cost of averting a new case of homelessness among very low-income families would be \$6800, 35% less than the per-person cost among all eligible callers (13–15).

The benefits of this intervention result from avoiding a number of costs associated with a person becoming homeless, including the cost of providing shelter and other housing services,

the cost to society of addressing other needs that may arise as a result of homelessness, and other private costs to the individual. One study estimates the average per-person cost of providing shelter for individuals who become homeless for the first time at about \$2400 in 2012 dollars (16), and another study estimates the per-person public costs (including health care, police and incarceration, and welfare programs such as food stamps) of the homeless to be \$5148 annually (17, 18).

Perhaps the most substantial benefits of homelessness prevention stem from reduced

Table 1. Mean characteristics and differences in means between treatment and control groups. Shown are the mean characteristics for the control groups and the regression-adjusted difference between groups. To test balance, each characteristic was regressed on an availability-of-funds indicator, fund-specific restriction variables, and seasonal control variables, and standard errors were clustered at the ZIP code level. Caller characteristics come from HPCC records from 20 January 2010 to 4 December 2012 for eligible, first-time callers; ZIP code-level characteristics come from the 2010–2012 American Community Surveys and are standardized to have a mean of 0 and a variance of 1. Shelter inhabitancy in the past 18 months comes from HMIS data for homeless shelters in Chicago.

Dependent variable	Homogeneous subsample		Main sample	
	Control group mean	Coefficient on availability of funds	Control group mean	Coefficient on availability of funds
Female	0.874	−0.100**	0.794	−0.032*
White, non-Hispanic	0.066	0.022	0.062	0.011
Black, non-Hispanic	0.906	−0.030	0.908	−0.017
Other, non-Hispanic	0.033	0.002	0.032	0.005
Hispanic	0.083	0.001	0.069	−0.001
Age	39.258	−1.078	39.129	−0.838*
Number of adults in caller's household	1.341	0.038	1.353	−0.019
Number of minors in caller's household	1.357	−0.012	1.389	−0.083
Standardized percentage in ZIP code with high school degree	−0.061	0.001	0.014	−0.034
Standardized labor force participation rate in ZIP code	−0.060	0.032	−0.015	−0.003
Standardized unemployment rate in ZIP code	0.017	0.005	0.014	0.005
Standardized median age in ZIP code	−0.163	0.150**	−0.006	0.018
Standardized monthly housing cost in ZIP code (in thousands)	0.034	−0.080	0.009	−0.040
Standardized median household income in ZIP code (in thousands)	−0.071	0.026	−0.004	−0.021
Standardized fraction black in ZIP code	0.046	−0.048	0.024	−0.012
Standardized fraction white in ZIP code	−0.042	0.027	−0.014	−0.010
Standardized fraction other races in ZIP code	−0.040	0.076	−0.037	0.051
Applying because of benefit loss	0.132	0.023	0.115	−0.010
Applying because of inability to pay bills	0.022	−0.015	0.013	−0.008**
Applying because exiting shared housing	0.044	0.013	0.089	0.012
Applying to flee abuse	0.016	−0.012	0.016	−0.002
Applying because of job loss	0.407	−0.030	0.301	0.012
Monthly income (in thousands)	1.175	0.080**	1.210	−0.052**
Receiving Supplemental Nutrition Assistance Program benefits	0.637	−0.057	0.633	−0.019
Receiving child support	0.061	−0.010	0.060	−0.017**
Receiving earned income	0.682	0.008	0.650	−0.004
Receiving disability payments	0.084	−0.026	0.096	−0.013
Receiving social security income	0.106	−0.003	0.137	0.000
Receiving income from Temporary Assistance for Needy Families	0.039	−0.001	0.061	0.003
Receiving unemployment payments	0.201	0.029	0.151	0.023
Receiving other income sources	0.067	−0.023	0.073	−0.005
Living situation: rent housing	0.907	−0.010	0.818	−0.015
Living situation: shared housing	0.093	0.010	0.174	0.015
Shelter inhabitancy in the past 18 months	0.027	0.003	0.029	0.008
<i>n</i>	182	1431	1858	4448

P* < 0.10; *P* < 0.05; two-tailed *t* test of the difference between treatment and control groups.

Table 2. ITT effects of fund availability on shelter spells. We performed regressions of shelter entry (3 or 6 months after calling the HPCC) or number of days spent in a shelter (6 months after calling the HPCC) on an availability-of-funds indicator, demographic characteristics of the caller, ZIP code-level characteristics, fund-specific restriction variables, and seasonal control variables. The standard errors were clustered at the ZIP code level. Shown here are estimates of the coefficients on the availability-of-funds indicator. Estimates of the coefficients on other covariates are given in table S8.

Dependent variable	Shelter admittance		Days spent in shelter
	3 months after calling	6 months after calling	6 months after calling
Main sample			
Funds are available	−0.014**	−0.016**	−2.620**
(Standard error)	(0.005)	(0.005)	(0.878)
<i>n</i>	4448	4448	4448
Mean of dependent variable for control group	0.016	0.021	3.132
Homogeneous subsample			
Funds are available	−0.019	−0.021	−2.987
(Standard error)	(0.013)	(0.013)	(1.911)
<i>n</i>	1431	1431	1431
Mean of dependent variable for control group	0.027	0.033	4.714
Rent callers			
Funds are available	−0.014**	−0.015**	−2.278**
(Standard error)	(0.006)	(0.006)	(1.025)
<i>n</i>	3574	3574	3574
Mean of dependent variable for control group	0.018	0.021	3.169
Security deposit callers			
Funds are available	−0.014*	−0.026**	−4.571**
(Standard error)	(0.008)	(0.010)	(1.538)
<i>n</i>	874	874	874
Mean of dependent variable for control group	0.012	0.020	3.062
Rent callers, below median income			
Funds are available	−0.023**	−0.022**	−3.389**
(Standard error)	(0.008)	(0.009)	(1.486)
<i>n</i>	1781	1781	1781
Mean of dependent variable for control group	0.023	0.025	3.565
Rent callers, above median income			
Funds are available	−0.004	−0.007	−0.942
(Standard error)	(0.008)	(0.009)	(1.352)
<i>n</i>	1790	1790	1790
Mean of dependent variable for control group	0.013	0.018	2.783

* $P < 0.10$; ** $P < 0.05$; two-tailed t test of the difference between treatment and control groups.

private costs. For example, homeless adults face higher mortality rates (19); one-fifth of homeless children have been separated from their families, and they are twice as likely to have a learning disability and to repeat a grade (20). These private costs of homelessness can be very high, so even a small reduction in homelessness can generate substantial cost savings. For example, the benefits associated with declined mortality alone can offset much of the cost of reducing homelessness through financial assistance. Estimates suggest that age-adjusted mortality rates for the homeless population in New York City are four times as great as for the U.S. population as a whole and two to three times as great as for the population of New York City (19, 21). Taking the lower bound of this mortality effect as causal, U.S. Environmental Protection Agency estimates of the value of a statistical life suggest that the

mortality-reducing benefit per person that avoids homelessness comes to about \$13,000 (supplementary materials, section 2.4). Based on this estimate, the benefits stemming from reduced mortality alone exceed the costs.

It is important to qualify that none of the estimates of the impact of homelessness on mortality are causal. In fact, there is virtually no evidence of the causal relationship between homelessness and other private costs. This makes it difficult to calculate reliable estimates of the private benefits stemming from homelessness prevention. Nevertheless, taking the estimates of potential benefits discussed above at face value, they amount to \$20,548 per homeless spell averted (\$2400 + \$5148 + \$13,000), which far exceeds the estimated \$10,300 price tag for reducing homelessness through emergency financial assistance. The implied net benefits could be even

greater if one were to include the value of other potential benefits of homelessness prevention, such as improved health, better academic outcomes for children, and others. Moreover, because the cost per homeless spell averted is much lower for the lowest-income callers, the net benefits of emergency assistance are considerably larger for this population. In addition, this discussion only addresses the impact of financial assistance on homelessness. Such assistance may also reduce the likelihood of moving involuntarily to another residence or moving in with family or friends. If temporary financial assistance reduces the likelihood of having to move, then the value of the benefits of this intervention would be even greater.

This study provides quasi-experimental evidence of the efficacy of a homelessness prevention approach that is widespread but understudied. The results show that temporary financial assistance does, in fact, prevent some individuals from having to enter a shelter. Although the cost of reducing homelessness through this prevention approach is relatively high, the cost savings of the program are likely to far exceed these costs because homelessness exacts a considerable toll on society and the homeless themselves. In addition, the evidence indicates that these programs would be more cost-effective if they were to target groups that benefit more from financial assistance, such as very low-income individuals and families.

A substantial expansion of temporary financial assistance programs may adversely affect the behavior of those who are potentially eligible—for example, by encouraging more individuals to seek assistance or by discouraging individuals from self-insuring through personal savings. Policymakers should take into account the potential for these sorts of moral-hazard responses when deciding whether to expand homelessness prevention programs.

REFERENCES AND NOTES

1. U.S. Department of Housing and Urban Development (HUD), *The 2014 Annual Homeless Assessment Report (AHAR) to Congress, Part I: Point-in-Time Estimates of Homelessness* (HUD Office of Community Planning and Development, 2014).
2. 211.org, “Find your local 2-1-1 service,” 2015; www.211us.org.
3. D. H. Friedman, J. Raymond, K. Puhala, T. Meschede, J. Tripp, M. Kala, *Preventing Homelessness and Promoting Housing Stability: A Comparative Analysis* (The Boston Foundation, 2007).
4. M. R. Burt, C. L. Pearson, A. E. Montgomery, *Strategies for Preventing Homelessness* (HUD Office of Policy Development and Research, 2005).
5. To conduct a joint test of the significance of these characteristics, we regressed availability of funds on all of the exogenous characteristics in Table 1, as well as on the fund-specific restrictions and seasonal controls. The F statistic for the null hypothesis that all coefficients on the exogenous variables are jointly equal to zero is $F_{(31, 40991)} = 1.36$, with $P = 0.09$.
6. This distinction is important because of noncompliance: Some callers who are referred to an agency for assistance do not receive funds. Furthermore, some callers seeking assistance when funds are not available may receive funds by calling back when funds are available. Unfortunately, our data sources do not include information on actual receipt of financial assistance. However, receipt information for a subset of HPCC callers suggests that just under three-quarters of those

- referred to funds actually receive financial assistance (7). Section 1.5 of the supplementary materials includes further discussion.
- C. George, J. Hilvers, K. Patel, D. Guelespe, *Evaluation of the Homelessness Prevention Call Center* (Loyola University Chicago Center for Urban Research and Learning, 2011).
 - These estimates reflect the effect of fund availability on the likelihood of checking into a shelter that reports data to the HMIS system (our source for shelter data). As explained in the supplementary materials, section 1.1, our shelter data do not cover all shelter entries. In addition, our data do not capture the effect for the unsheltered homeless or for those who move in with others.
 - Because these subgroup analyses entail multiple hypotheses, there is greater risk of a type I error. If we make a simple Bonferroni adjustment to the *P* values, treating each sample split in isolation (i.e., the number of subgroups is two), then two of the main effects for the subgroups in table S10 are significant at *P* < 0.05: the effect for those below median income and the effect for those calling in the winter. If, however, one is interested in controlling for the familywise error rate for all six regressions in table S10 simultaneously, then none of the main effects are significant at *P* < 0.05, but the effect for those below median income is significant at *P* < 0.1 (*P* = 0.06).
 - M. Shinn, J. Baumohl, K. Hopper, *Anal. Soc. Issues Public Policy* **1**, 95 (2001).
 - D. P. Culhane, S. Metraux, T. Byrne, *Housing Policy Debate* **21**, 2 (2010).
 - Only about 2% of those in our main sample who called when no funds were available entered a shelter within the next 6 months. We estimate that our data capture just over half of all homeless spells (supplementary materials, section 1.1), suggesting that about 4% of these callers become homeless. However, financial assistance may also reduce the risk of other bad outcomes such as involuntarily having to move to another residence or having to move in with family or friends.
 - Other studies have noted that homelessness prevention programs could be more effective if they were to do a better job of targeting resources to high-risk families and individuals (14, 15).
 - A. L. Greer, M. Shinn, J. Kwon, S. Zuiderveen, *Soc. Serv. Rev.* **90**, 130–155 (2016).
 - M. Shinn, A. L. Greer, J. Bainbridge, J. Kwon, S. Zuiderveen, *Am. J. Public Health* **103**, S324–S330 (2013).
 - B. Spellman, J. Khadduri, B. Sokol, J. Leopold, Abt Associates, *Costs Associated with First-Time Homelessness for Families and Individuals* (HUD Office of Policy Development and Research, 2010).
 - D. Flaming, H. Toros, P. Burns, *Home Not Found: The Cost of Homelessness in Silicon Valley* (Economic Roundtable, 2015).
 - These are very rough estimates. The estimate of public costs is for Los Angeles only and is for the absolute per-person cost for the homeless, rather than the cost relative to an otherwise similar person who is not homeless. But this estimate might be a reasonable upper bound of the impact of homelessness on the costs for health care, police and incarceration, and welfare programs.
 - S. M. Barrow, D. B. Herman, P. Córdova, E. L. Struening, *Am. J. Public Health* **89**, 529–534 (1999).
 - National Alliance to End Homelessness, “Rapid re-housing: Successfully ending family homelessness,” 21 May 2012; www.endhomelessness.org/library/entry/rapid-re-housing-successfully-ending-family-homelessness.
 - The estimates from this study for the effect of homelessness on mortality are for a cross section of single adults in New York City. The effect of homelessness on mortality for this group may not accurately reflect that for the typical caller to the HPCC. Because there is very limited evidence on the direct effect of homelessness on mortality, these are the best estimates available.

ACKNOWLEDGMENTS

This project is supported by the Wilson Sheehan Lab for Economic Opportunities at the University of Notre Dame. The study received approval from the University of Notre Dame's Institutional Review Board (protocol 14-10-2111). We have benefited from the research assistance of A. Kroeger, J. Macke, and C. Palmer, as well as from the help and cooperation of Catholic Charities of the Archdiocese of Chicago and its HPCC, with special thanks to K. Donahue,

B. Haennicke, S. Murray, and N. Russo. We also thank N. Brennan, P. Thangaraj, and the data team at All Chicago, who provided the HMIS data. We thank K. Corinth, S. Heller, and seminar participants at the University of Wisconsin Institute for Research on Poverty, the University of Chicago, the University of Maryland, the University of Utah, and the University of Notre Dame for helpful comments. Data and code to replicate the results of this study are available at the University of Michigan's Interuniversity Consortium for Political and Social Research (www.openicpsr.org/repoEntry/show/83434).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/694/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S7
Tables S1 to S10
References (22–28)

6 May 2016; accepted 19 July 2016
10.1126/science.aag0833

ATMOSPHERIC CHEMISTRY

Atmospheric photochemistry at a fatty acid-coated air-water interface

Stéphanie Rossignol,^{1*} Liselotte Tinel,^{1*} Angelica Bianco,² Monica Passananti,¹ Marcello Brigante,² D. James Donaldson,^{3,4†} Christian George^{1†}

Although fatty acids are believed to be photochemically inert in the actinic region, complex volatile organic compounds are produced during illumination of an air-water interface coated solely with a monolayer of carboxylic acid. When aqueous solutions containing nonanoic acid (NA) at bulk concentrations that give rise to just over a monolayer of NA coverage are illuminated with actinic radiation, saturated and unsaturated aldehydes are seen in the gas phase, and more highly oxygenated products appear in the aqueous phase. This chemistry is probably initiated by triplet-state NA molecules excited by direct absorption of actinic light at the water surface. Because fatty acids-covered interfaces are ubiquitous in the environment, such photochemical processing will have a substantial impact on local ozone and particle formation.

Over the past 20 years, interfacial processes have become increasingly of interest in the field of atmospheric chemistry (1), with many studies showing that environmental surfaces display specific chemistry and photochemistry, enhancing certain reactions and acting as reactive sinks or sources for various atmospherically relevant species (2–5). Many molecules display a free-energy minimum at the air-water interface (6–9), making it a favored venue for compound accumulation and reaction. Indeed, surface-active molecules have been shown to undergo specific photochemistry at the air-water interface; for example, dimers of 2-oxooctanoic acid are formed there in addition to the expected fragmentation pathways (10). Of importance in the environment, the sea-surface microlayer (SML) is mainly composed of surface-active, biogenically derived organics. Recent measurements of the gas phase above irradiated SML surfaces, or simple organic-coated aqueous samples con-

taining natural photosensitizers, have revealed the photochemical formation of a wide variety of functionalized volatile organic compounds (VOCs) (11, 12). Such previously unknown surface-photosensitized chemistry could constitute a substantial abiotic source of VOCs to the marine boundary layer; it is therefore important to fully understand the production mechanisms. Here, we show that complex VOCs may also be produced without an added photosensitizer from an illuminated simple organic acid, which does not absorb actinic light in dilute solutions but exhibits a totally different behavior once present as coating at the air-water interface.

Nonanoic acid (NA) is a highly surface-active simple organic fatty acid, representative of many biogenic compounds present in the SML. In the gas phase and at its solubility limit in aqueous solution, it does not absorb ultraviolet (UV)-visible radiation in the actinic region (that is, at wavelengths longer than ~280 nm). Surprisingly, when a 15-mL quartz reactor was filled halfway with a 2 mM aqueous solution of NA, a concentration giving rise to just over a monolayer, and irradiated for 1 hour with a Xenon arc lamp, prompt formation of gas-phase saturated and unsaturated C₉ and C₈ aldehydes was observed. This result is displayed in Fig. 1. In addition to the aldehydes, a wide variety of (mostly oxygenated) photoproducts was detected in both gas and condensed phases; these are listed in table S1.

The influence of oxygen on the product distribution is displayed for the gas and solution

¹Université Lyon, Université Claude Bernard Lyon 1, CNRS, Institut de Recherches sur la Catalyse et l'Environnement de Lyon (IRCELYON), F-69626 Villeurbanne, France, 2 avenue Albert Einstein, F-69626 Villeurbanne, France. ²Clermont Université, Université Blaise Pascal, Institut de Chimie de Clermont-Ferrand, BP 10448, F-63000 Clermont-Ferrand, France. ³Department of Chemistry, University of Toronto, 80 St. George St., Toronto, ON M5S 3H6, Canada. ⁴Department of Physical and Environmental Sciences, University of Toronto, 80 St. George St., Toronto, ON M5S 3H6, Canada. *These authors contributed equally to this work. †Corresponding author. Email: christian.george@ircelyon.univ-lyon1.fr (C.G.); jdonalds@chem.utoronto.ca (D.J.D.)

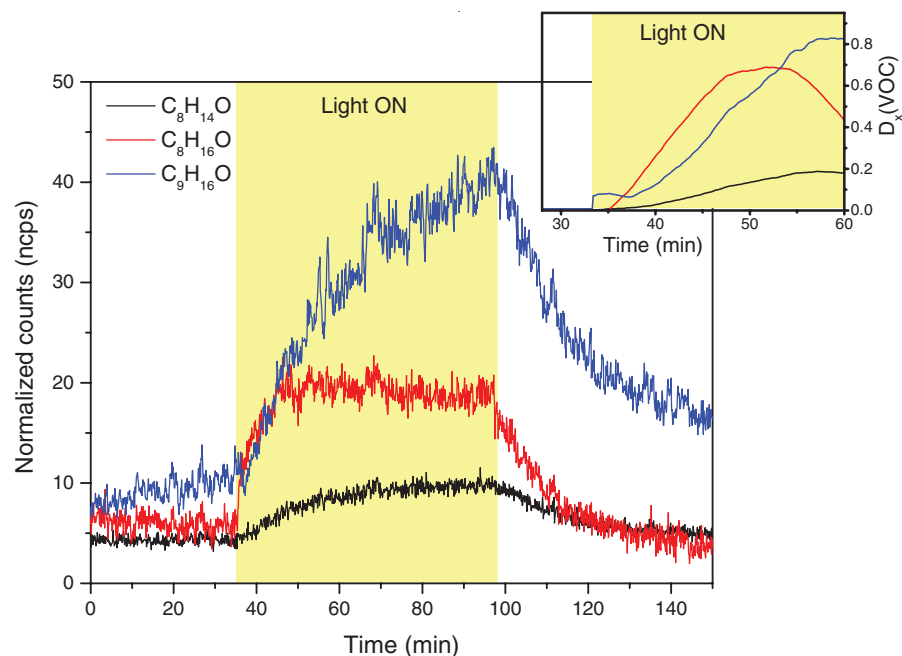


Fig. 1. Mass spectrometry time traces using NO^+ ionization of the C_9 -unsaturated, C_8 -saturated, and C_8 -unsaturated aldehydes observed in the gas phase during illumination of the NA-coated water surface. Blue, C_9 -unsaturated; red, C_8 -saturated; black, C_8 -unsaturated. (Inset) The first-order derivatives of the first 28 min of irradiation, showing the different slopes for each compound (supplementary materials).

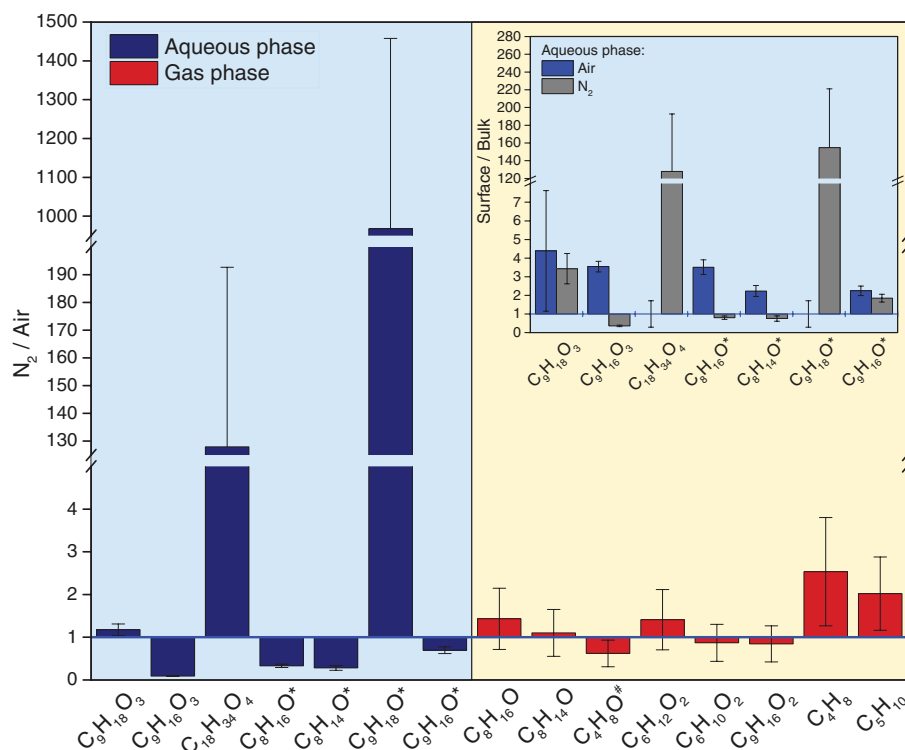


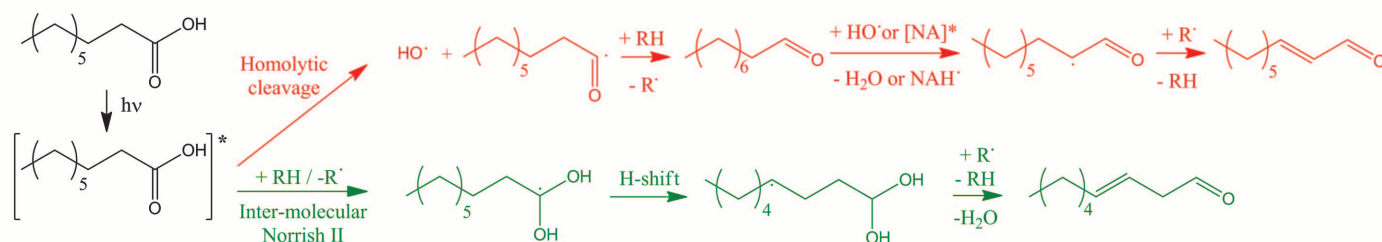
Fig. 2. Comparison of the enrichment of the net production of selected compounds in the condensed phase and the gas phase under N_2 compared with the production in air. (Inset) The enrichment of the net production in the condensed phase when an irradiated surface is present compared with a bulk irradiated-only experiment, under air and N_2 . Errors reflect the analytical uncertainties (condensed phase) or repeatability between experiments (gas phase). Asterisk indicates detected as PFBHA derivatives. Hash mark indicates identified as ketone owing to NO^+ adduct.

phase in Fig. 2, by comparing the ratios of net production of some of the observed products obtained in an atmosphere of N_2 versus under air. Saturated and unsaturated aldehydes and ketones ($\leq \text{C}_9$) are observed in both phases. In the condensed phase, C_9 products with three or more oxygen atoms, tentatively identified as NA bearing additional hydroxyl and/or carbonyl functions, were also detected. These are clearly favored when oxygen is more abundant. On the contrary, alkenes, such as C_4H_8 and C_5H_{10} , were detected in the gas phase and were favored in the absence of oxygen. Additionally, $\text{C}_{18}\text{H}_{34}\text{O}_4$, the covalently bonded dimer of the parent acid, was observed in the condensed phase and favored in an oxygen-poor environment.

The surface-specific nature of the chemistry is further illustrated in the Fig. 2 inset. For some experiments, the quartz reactor was filled completely so as to eliminate the air-water interface exposed to light. The enrichment observed in the experiment with an irradiated surface (defined as the ratio of the net production with irradiated surface over the production without irradiated surface) clearly demonstrates that both the $\text{C}_{18}\text{H}_{34}\text{O}_4$ product and the C_9 saturated monocarbonyl compound ($\text{C}_9\text{H}_{16}\text{O}$) are formed only when there is an interface present in the absence of oxygen. This result also shows that reactions taking place on the quartz surfaces of the reactor are not responsible for the observations.

The results outlined above are consistent with an interface-specific radical-initiated reaction, with a mechanism similar to those we have previously reported for reactions initiated by a known photosensitizer (11–13). The notable and surprising finding here is that there is no added photosensitizing agent to initiate the chemistry. Of course, this immediately suggests that an impurity is responsible. High-performance liquid chromatography–mass spectrometry (HPLC-MS) analysis of the NA samples gave no detectable aldehyde impurities. When a higher-purity grade (99.5%) of NA was used, no notable difference in the products was observed upon irradiation (fig. S1). Likewise, inductively coupled plasma (ICP) analysis of the 97% purity NA (described in the supplementary materials) shows that photoactive transition metals—in particular, Fe and Cu—are not responsible for initiating this chemistry. The experiments were repeated by using a Pyrex filter with a 50% transmittance cut-off of 315 nm. Although the total amount of product formation after 1 hour of irradiation was smaller than without the filter, by extending the irradiation time to 5 hours, most of the previously detected products were observed with intensities comparable with those obtained after 1 hour without the optical filter (fig. S2). These tests lead us to conclude that NA is somehow itself responsible for the observed photochemistry.

Such photochemistry has not been reported previously, and dilute aqueous solutions of NA do not absorb in the near-UV range. However, the UV-visible absorption spectra of neat or concentrated solutions of NA reveal an interesting feature. In addition to the main absorption band

Initiation and formation of C₉ aldehydes:

HO• or [NA]* can react with NA according to:

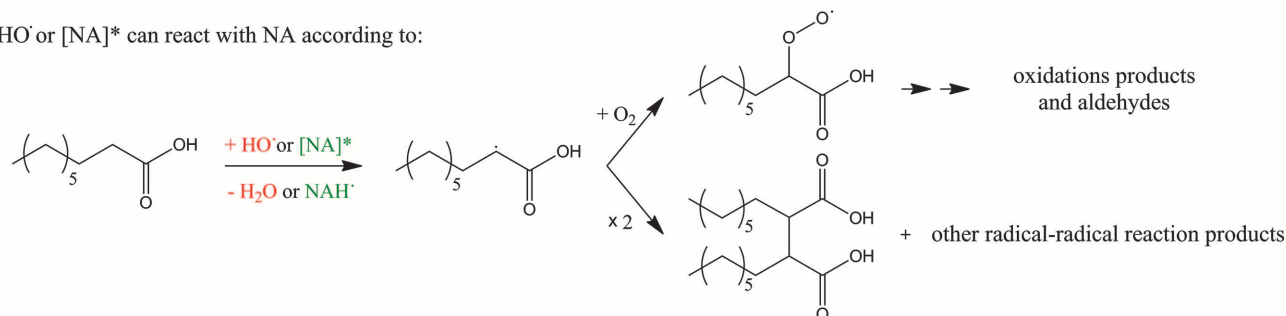


Fig. 3. Reaction mechanisms for NA after its excitation to the triplet state.

of NA, centered as expected at 212 nm (14), a second very weak absorption band is seen (fig. S3). This is centered at 270 nm and extends to ~330 nm, growing linearly in intensity with increasing NA. Such weak absorption bands centered around 270 nm have been observed previously for a series of neat small aliphatic carboxylic acids (15). Although the acid anhydride is expected to absorb in this region (16), it seems highly unlikely that this species would be found in aqueous solution. If instead there is a very weakly absorbing reactive state of the acid, favored at the interface, its existence could explain the initiation of photochemistry seen here. The overlap between the absorption of neat NA and the solar spectrum at Earth's surface is illustrated in fig. S4.

Indeed, quantum chemical calculations (described fully in the supplementary materials), carried out on propionic acid using density-functional theory at the B3LYP 6-311+G(2df, 2p) level, indicate the presence of a triplet state lying at 402 kJ/mol above the ground state, whose Franck-Condon maximum is predicted to be at 443 kJ/mol (270 nm). Although direct absorption to a triplet is expected to be weak, in the present 2 mM NA experiments the acid is quite concentrated at the water surface, as indicated in the adsorption isotherm displayed in fig. S5. With a monolayer coating indicated to form near 0.6 mM, the 2 mM solutions will exhibit multilayers of NA at the surface, perhaps giving rise to light absorption similar to that given by a neat solution of NA.

If a triplet state is responsible for initiating the surface photochemistry, one can imagine several possible initiation steps. The direct photochemistry of carboxylic acids in the 200- to 260-nm range can lead to a photolytic cleavage via a Norrish type I reaction mechanism, giving rise to the formation of radicals. If the triplet state

reacts similarly, one might expect the products to include octyl radicals, but this cannot explain the formation of the observed dimers of NA or any of the observed C₉ compounds. Another possibility is the direct dissociation pathway that forms OH and nonyl radicals. Our quantum chemical calculations indicate that this channel is energetically possible for the triplet, being endothermic by 437 kJ/mol for the ground-state species, so near thermoneutral for the triplet. If OH is formed, it is expected to give a suite of products similar to what is observed, as shown by experiments using H₂O₂ as a photoinitiator and as displayed in fig. S6. The acyl radical could abstract a hydrogen from an adjacent acid to form the C₉-saturated aldehyde and an acid radical; this reaction is predicted to be only weakly endothermic (~6 kJ/mol) for the C₃ system and so may represent a formation route to the C₉-saturated aldehyde.

An alternative (or parallel) reaction initiation could involve the excited triplet acid acting as a “traditional” triplet photosensitizer: abstracting a hydrogen from an adjacent acid molecule to form a diol radical and an acid radical. This “intermolecular Norrish type II” reaction is predicted to be energetically feasible (exothermic by ~61 kJ/mol), with a low transition-state energy (supplementary materials). The formation of a diol radical may explain the formation of both the C₉ carbonyls, nonanal and nonenal, via diol dehydration in the water-poor organic surface layer.

Both initiation mechanisms appear to compete with classical hydrogen abstraction initiated by OH radicals. Indeed, the experiments performed in the presence of H₂O₂ in the aqueous phase also show the formation of C₉ carbonyls (fig. S6), whose formation can only be explained by these newly proposed chemistries. A general mechanism for the NA photosensitized degradation

follows that proposed for octanol at the air-water interface (13). As illustrated in Fig. 3, after H-abstraction from the acid (by an excited triplet acid, OH, or an organic radical), two fates for the resulting C₉ acid radical are expected: the addition of molecular oxygen and radical-radical recombination. Oxygen addition, followed by disproportionation and decomposition, is expected to be favored in the presence of O₂. This pathway is in competition with radical recombination and other radical-radical reactions, which are expected if there is a high concentration of radicals or the availability of O₂ is reduced. The latter pathway readily explains the greatly enhanced appearance of NA dimer recombination products in the N₂-only experiments. This type of radical recombination, which is fast, has been observed previously in photosensitized reactions in aerosols (17) and at the air-water surface (9, 12) but occurs only where high radical concentrations are achieved (18). The presence of such recombination products highlights some features of the surface-active enhanced concentration at the air-water interface, which could also then favor disproportionation reactions (19) and explain the formation of nonenoic acid.

We report the photochemical production of functionalized and unsaturated compounds at a NA-coated air-water interface in the absence of any known photosensitizer. This photochemistry may indicate the involvement of a triplet state common to all carboxylic acid molecules, accessible after light absorption in the environmentally relevant 280- to 330-nm UV region. The weak absorption to this state suggests that accessing the triplet is only possible for concentrated solutions of NA, such as exist for multilayers of surface-active acids at the air-water interface, and could be particularly relevant for longer-chain acids (20). Reaction could be initiated by

dissociation of this excited state, producing radicals, or by the formation of a diol radical after reaction of an excited-state fatty acid with an adjacent molecule.

Because fatty acid-covered surfaces are ubiquitous, the photochemical production of gas-phase unsaturated and functionalized compounds will affect the local oxidative capacity of the atmosphere and will lead to secondary aerosol formation. This interfacial photochemistry may exert a very large impact, especially if in general the mere presence of a surface layer of a carboxylic acid can trigger this interfacial photochemistry at ocean surfaces, cloud droplets, and the surface of evanescent aerosol particles.

REFERENCES AND NOTES

- C. George, M. Ammann, B. D'Anna, D. J. Donaldson, S. A. Nizkorodov, *Chem. Rev.* **115**, 4218–4258 (2015).
- A. M. Baergen, D. J. Donaldson, *Environ. Sci. Technol.* **47**, 815–820 (2013).
- Y. Dupart *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20842–20847 (2012).
- C. Zhu, B. Xiang, L. Zhu, R. Cole, *Chem. Phys. Lett.* **458**, 373–377 (2008).
- S. Enami, M. R. Hoffmann, A. J. Colussi, *J. Phys. Chem. Lett.* **6**, 527–534 (2015).
- M. T. C. Martins-Costa, J. M. Anglada, J. S. Francisco, M. F. Ruiz-Lopez, *J. Am. Chem. Soc.* **134**, 11821–11827 (2012).
- M. T. C. Martins-Costa, F. F. Garcia-Prieto, M. F. Ruiz-Lopez, *Org. Biomol. Chem.* **13**, 1673–1679 (2015).
- K. Mozgawa, B. Mennucci, L. Frediani, *J. Phys. Chem. C* **118**, 4715–4725 (2014).
- R. Vácha, P. Slaviček, M. Mucha, B. J. Finlayson-Pitts, P. Jungwirth, *J. Phys. Chem. A* **108**, 11573–11579 (2004).
- E. C. Griffith, R. J. Rapf, R. K. Shoemaker, B. K. Carpenter, V. Vaida, *J. Am. Chem. Soc.* **136**, 3784–3787 (2014).
- R. Ciuraru *et al.*, *Environ. Sci. Technol.* **49**, 13199–13205 (2015).
- R. Ciuraru *et al.*, *Sci. Rep.* **5**, 12741 (2015).
- H. Fu *et al.*, *J. Am. Chem. Soc.* **137**, 8348–8351 (2015).
- J. D. Coyle, *Chem. Rev.* **78**, 97–123 (1978).
- L. R. Caswell, M. F. Howard, T. M. Onisto, *J. Org. Chem.* **41**, 3312–3316 (1976).
- J. G. Calvert, J. N. Pitts, *Photochemistry* (Wiley, 1966).
- S. Rossignol *et al.*, *Environ. Sci. Technol.* **48**, 3218–3227 (2014).
- A. F. Parsons, *An Introduction to Free Radical Chemistry* (Blackwell Science, 2000).
- J. M. C. Plane *et al.*, "Photochemistry in the sea-surface microlayer" in *The Sea Surface and Global Change*, P. S. Liss, R. A. Duce, Eds. (Cambridge Univ. Press, 1997), pp. 71–93.
- J. R. Kanicky, A. F. Poniatowski, N. R. Mehta, D. O. Shah, *Langmuir* **16**, 172–177 (2000).

ACKNOWLEDGMENTS

This study was supported by the European Research Council (ERC) under the European Union's Seventh Framework Program (FP/2007-2013)/ERC Grant Agreement 290852-AIRSEA. D.J.D. acknowledges ongoing support from the Natural Sciences and Engineering Research Council of Canada. The authors are grateful to P. Mascunan and N. Cristin for the ICP-MS analysis and N. Charbonnel and S. Perrier for the technical support provided by IRCELYON. All the data presented here can be downloaded from the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/699/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S3
References (21–26)
Database S1

29 January 2016; accepted 23 June 2016
10.1126/science.aaf3617

LIFE HISTORY

Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (*Somniosus microcephalus*)

Julius Nielsen,^{1,2,3,4,*} Rasmus B. Hedeolm,² Jan Heinemeier,⁵ Peter G. Bushnell,⁶ Jørgen S. Christiansen,⁴ Jesper Olsen,⁵ Christopher Bronk Ramsey,⁷ Richard W. Brill,^{8,9} Malene Simon,¹⁰ Kirstine F. Steffensen,¹ John F. Steffensen¹

The Greenland shark (*Somniosus microcephalus*), an iconic species of the Arctic Seas, grows slowly and reaches >500 centimeters (cm) in total length, suggesting a life span well beyond those of other vertebrates. Radiocarbon dating of eye lens nuclei from 28 female Greenland sharks (81 to 502 cm in total length) revealed a life span of at least 272 years. Only the smallest sharks (220 cm or less) showed signs of the radiocarbon bomb pulse, a time marker of the early 1960s. The age ranges of prebomb sharks (reported as midpoint and extent of the 95.4% probability range) revealed the age at sexual maturity to be at least 156 ± 22 years, and the largest animal (502 cm) to be 392 ± 120 years old. Our results show that the Greenland shark is the longest-lived vertebrate known, and they raise concerns about species conservation.

The Greenland shark (Squaliformes, *Somniosus microcephalus*) is widely distributed in the North Atlantic, with a vertical distribution ranging from the surface to at least 1816-m depth (1, 2). Females outgrow males, and adults typically measure 400 to 500 cm, making this shark species the largest fish native to arctic waters. Because reported annual growth is ≤1 cm (3), their longevity is likely to be exceptional. In general, the biology of the Greenland shark is poorly understood, and longevity and age at first reproduction are completely unknown. The species is categorized as "Data Deficient" in the Norwegian Red List (4).

Conventional growth zone chronologies cannot be used to age Greenland sharks because of their lack of calcified tissues (5). To circumvent this problem, we estimated the age from a chronology obtained from eye lens nuclei by applying radiocarbon dating techniques. In vertebrates,

the eye lens nucleus is composed of metabolically inert crystalline proteins, which in the center (i.e., the embryonic nucleus) is formed during prenatal development (6, 7). This tissue retains proteins synthesized at approximately age 0: a unique feature of the eye lens that has been exploited for other difficult-to-age vertebrates (6, 8, 9).

Our shark chronology was constructed from measurements of isotopes in the eye lens nuclei from 28 female specimens (81 to 502 cm total length, table S1) collected during scientific surveys in Greenland during 2010–2013 (fig. S1) (see supplementary materials). We used radiocarbon (¹⁴C) levels [reported as percent of modern carbon (pMC)] to estimate ages and stable isotopes, ¹³C and ¹⁵N (table S1), to evaluate the carbon source (supplementary materials). Depleted ^δ¹³C and enriched ^δ¹⁵N levels established that the embryonic nucleus radiocarbon source was of dietary origin and represents a high trophic level. In other words, isotope signatures are dictated by the diet of the shark's mother, not the sampled animals. We set the terminal date for our analyses to 2012, because samples were collected over a 3-year period. The chronology presumes that size and age are positively correlated.

Since the mid-1950s, bomb-produced radiocarbon from atmospheric tests of thermonuclear weapons has been assimilated in the marine environment, creating a distinct "bomb pulse" in carbon-based chronologies (10). The period of rapid radiocarbon increase is a well-established time stamp for age validation of marine animals (11–14). Radiocarbon chronologies of dietary origin (reflecting the food web) and chronologies reflecting dissolved inorganic radiocarbon of surface mixed and deeper waters, have shown that the timing of the bomb pulse onset (i.e., when

¹Marine Biological Section, University of Copenhagen, Strandpromenaden 5, 3000 Helsingør, Denmark. ²Greenland Institute of Natural Resources, Post Office Box 570, Kivioq 2, 3900 Nuuk, Greenland. ³Den Blå Planet, National Aquarium Denmark, Jacob Fortlingsvej 1, 2770 Kastrup, Denmark. ⁴Department of Arctic and Marine Biology, UiT The Arctic University of Norway, 9037 Tromsø, Norway. ⁵Aarhus AMS Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus, Denmark. ⁶Department of Biological Sciences, Indiana University South Bend, 1700 Mishawaka Avenue, South Bend, IN, USA. ⁷Oxford Radiocarbon Accelerator Unit, University of Oxford, Dyson Perrins Building, South Parks Road, Oxford OX1 3QY, UK. ⁸National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, James J. Howard Marine Sciences Laboratory, 74 Magruder Road, Highlands, NJ 07732, USA. ⁹Virginia Institute of Marine Science, Post Office Box 1346, Gloucester Point, VA 23062, USA. ¹⁰Greenland Climate Research Centre, Greenland Institute of Natural Resources, Post Office Box 570, Kivioq 2, 3900 Nuuk, Greenland. *Corresponding author. Email: julius.nielsen@bio.ku.dk

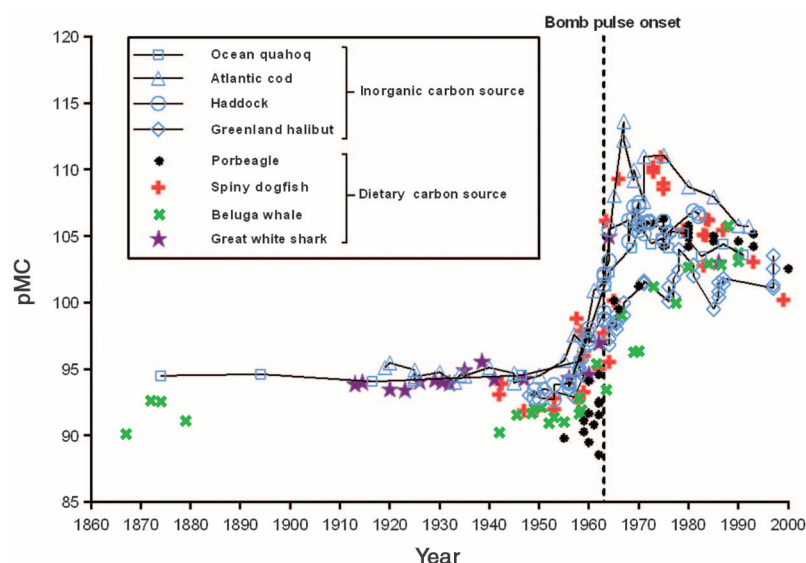


Fig. 1. Radiocarbon chronologies of the North Atlantic Ocean. Radiocarbon levels (pMC) of different origin (inorganic and dietary) over the past 150 years are shown. Open symbols (connected) reflect radiocarbon in marine carbonates (inorganic carbon source) of surface mixed and deeper waters (26, 36–38). Solid symbols reflect radiocarbon in biogenic archives of dietary origin (11, 14, 22, 24). The dashed vertical line indicates the bomb pulse onset in the marine food web in the early 1960s.

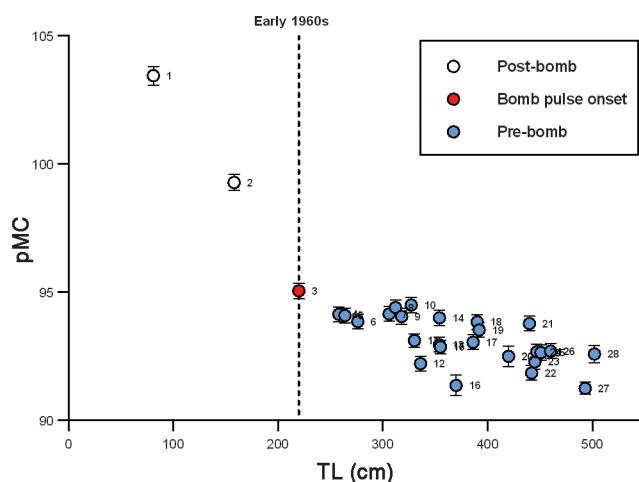


Fig. 2. Radiocarbon in eye lens nuclei of Greenland sharks. Radiocarbon levels (pMC \pm SD, table S1) from 28 females plotted against total length (TL) are shown. Individual animals are identified by the numbers next to the symbols. Nos. 1 and 2 are of postbomb origin, and nos. 4 to 28 are of prebomb origin. We consider shark no. 3 to be from the early 1960s, which is the latest timing of the bomb pulse onset (dashed vertical line).

bomb-produced radiocarbon becomes detectable in a chronology) is synchronous within a few years and no later than early 1960s across the northern North Atlantic (Fig. 1).

Sexually mature females >400 cm have been caught across the Greenland continental shelf at depths between 132 and ~ 1200 m [(15, 16) and table S1]. Their diet (15–17) and stable isotope signatures (18) (table S1) are comparable to those of other marine top predators such as the porbeagle (*Lamna nasus*), white shark

(*Carcharodon carcharias*), spiny dogfish (*Squalus acanthias*), and beluga whale (*Delphinapterus leucas*) (11, 14, 19–24), for which the bomb pulse onset has been established (Fig. 1). We therefore consider the early 1960s as appropriate for the timing of the bomb pulse onset for the Greenland shark chronology as well.

The two smallest animals (nos. 1 and 2) had the highest radiocarbon levels (>99 pMC), implying that they were indeed affected by the bomb pulse (Fig. 2). However, given the variability

of bomb pulse curves (Fig. 1), no exact age can be assigned to these animals other than that they were born later than the early 1960s. The third animal in the chronology (no. 3, 95.06 pMC), on the other hand, had a radiocarbon level slightly above those of the remaining sharks (nos. 4 to 28, pMC <95), placing its birth year close to the same time as the bomb pulse onset (i.e., early 1960s, Fig. 2). We therefore assign shark no. 3 (total length 220 cm) an age of ~ 50 years in 2012 and consider the remaining 25 larger animals to be of prebomb origin.

We estimated the age of prebomb sharks based on the Marine13 radiocarbon calibration curve (25), which evaluates carbon-based matter predating the bomb pulse that originates from surface mixed waters. The observed synchronicity of the bomb pulse onset (Fig. 1) supports the presumption that natural temporal changes of prebomb radiocarbon are imprinted in the marine food webs with negligible delay. We contend that the Marine13 curve can contribute to the assessment of the age of prebomb sharks despite the difficulties associated with (i) the low variation in the radiocarbon curve over the past 400 years (25); and (ii) that the degree of radiocarbon depletion in contemporaneous surface mixed waters (local reservoir age deviations, ΔR) differs between regions (26), meaning that the carbon source of the eye lens nucleus reflects food webs of potentially different ΔR levels. Consequently, radiocarbon levels of prebomb animals must be calibrated as a time series under a set of biological and environmental constraints.

We used OxCal (version 4.2) to do this calibration (27). The program uses Bayesian statistics to combine prior knowledge with calibrated age probability distributions to provide posterior age information (28, 29). We constrained age ranges with presumptions about von Bertalanffy growth, size at birth, the age of animal no. 3 deduced from the bomb pulse onset (biological constraints), and plausible ΔR levels from the recent past (environmental constraint). This makes up a Bayesian model that is detailed in the supplementary materials.

Calibrations of single pMC measurements without biological constraints are shown as probability distributions of age with very wide ranges (light blue distributions, Fig. 3). When imposing the model, constrained and narrower age estimates are produced for each prebomb individual, shown as posterior probability distributions of age (dark blue distributions) in Fig. 3 and posterior calibrated age ranges at 95.4% (2σ) probability in table S2. OxCal also calculated agreement indices for each individual shark (A index) and for the calibration model (A_{model}). This allowed us to evaluate the consistency between modeled age ranges and Marine13, as well as the internal agreement between data points of the model (table S2) (30). To test the effect of the fixed age parameter (shark no. 3), a sensitivity analysis was made (supplementary materials and fig. S2), showing that the overall finding of extreme Greenland shark longevity is robust

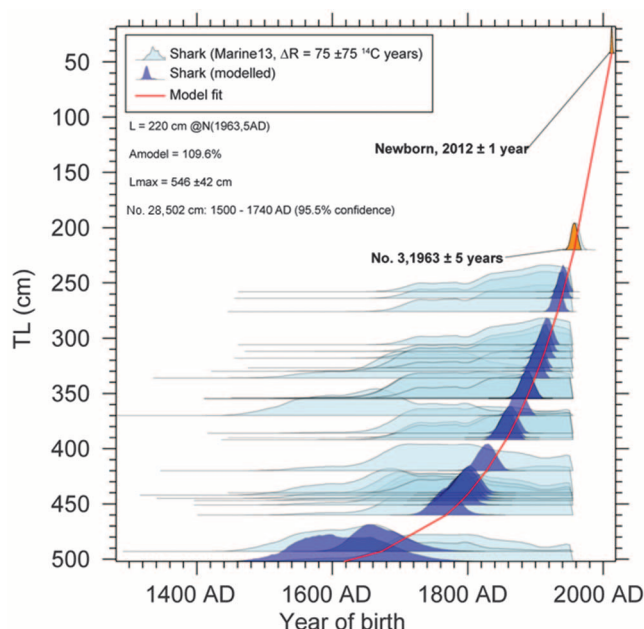


Fig. 3. Bayesian age ranges of prebomb sharks. The estimated year of birth against total length (TL) for prebomb sharks (nos. 4 to 28) is shown. Light blue shows the individual age probability distributions for each shark, and modeled posterior age probability distributions are shown in dark blue. Fixed age distributions (model input) of one newborn shark (42 cm, 2012 \pm 1) and of shark no. 3 (220 cm, born in 1963 \pm 5) are shown in orange. The red line is the model fit connecting the geometric mean for each posterior age probability distribution. (**Inset**) The model output; i.e., A_{model} , L_{max} , and range of birth year for shark no. 28. Also see the supplementary materials.

regardless of the exact timing of the bomb pulse onset (1958–1980).

The model estimated asymptotical total length to be 546 \pm 42 cm (mean \pm SD), a size matching the largest records for Greenland sharks (2), and the age estimates (reported as midpoint and extent of the 95.4% probability range) of the two largest Greenland sharks to be 335 \pm 75 years (no. 27, 493 cm) and 392 \pm 120 years (no. 28, 502 cm). Moreover, because females are reported to reach sexual maturity at lengths >400 cm (15), the corresponding age would be at least 156 \pm 22 years (no. 19, 392 cm) (table S2). A_{model} was 109.6%, demonstrating that samples are in good internal agreement, implying that the age estimates are reliable.

The validity of our Greenland shark age estimates is supported by other lines of evidence. For instance, we found sharks <300 cm to be younger than 100 years (table S2). Such age estimates are indirectly corroborated by their depleted $\delta^{13}\text{C}$ levels (table S1), possibly reflecting the Suess effect, another chemical time mark triggered by emissions of fossil fuels, imprinted in marine food webs since the early 20th century (31, 32). In addition, high levels of accumulated anthropogenic contaminants may suggest that ~300-cm females are older than 50 years (33). Taken together, these findings seem to corroborate an estimated life span of at least 272 years for Greenland sharks attaining more than 500 cm in length.

Our results demonstrate that the Greenland shark is among the longest-lived vertebrate spe-

cies, surpassing even the bowhead whale (*Balaena mysticetus*, estimated longevity of 211 years) (9). The life expectancy of the Greenland shark is exceeded only by that of the ocean quahog (*Arctica islandica*, 507 years) (34). Our estimates strongly suggest a precautionary approach to the conservation of the Greenland shark, because they are common bycatch in arctic and subarctic groundfish fisheries and have been subjected to several recent commercial exploitation initiatives (35).

REFERENCES AND NOTES

- H. B. Bigelow, W. C. Schroeder, in *Fishes of the Western North Atlantic*, A. E. Parr, Ed. (Yale University, New Haven, CT, 1948), pp. 516–523.
- S. E. Campana, A. T. Fisk, A. P. Klimley, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **115**, 109–115 (2015).
- P. M. Hansen, *International Commission for the Northwest Atlantic Fisheries Special Publication* **4**, 172–175 (1963).
- S. Henriksen, O. Hilmo, Eds., *Norsk Rødliste for Arter* (Artsdatabanken, Norge, 2015).
- P. M. Kyne, C. A. Simpendorfer, Adaptive physiology and conservation, in *Sharks and Their Relatives*, J. C. Carrier, J. A. Musick, M. R. Heithaus, Eds. (CRC Press, 2010), pp. 37–71.
- N. Lynnerup, H. Kjeldsen, S. Heegaard, C. Jacobsen, J. Heinemeier, *PLOS ONE* **3**, e1529 (2008).
- S. Bassnett, Y. Shi, G. F. J. M. Vrensen, *Philos. Trans. R. Soc. London Ser. B* **366**, 1250–1264 (2011).
- J. L. Bada, C. D. Vrolijk, S. Brown, E. R. M. Druffel, R. E. M. Hedges, *Geophys. Res. Lett.* **14**, 1065–1067 (1987).
- J. C. George et al., *Can. J. Zool.* **77**, 571–580 (1999).
- H. De Vries, *Science* **128**, 250–251 (1958).
- S. E. Campana, L. J. Natanson, S. Myklevoll, *Can. J. Fish. Aquat. Sci.* **59**, 450–455 (2002).
- J. M. Kalish, *Earth Planet. Sci. Lett.* **114**, 549–554 (1993).

- M. P. Francis, S. E. Campana, C. M. Jones, *Mar. Freshw. Res.* **58**, 10–23 (2007).
- L. L. Hamady, L. J. Natanson, G. B. Skomal, S. R. Thorrold, *PLOS ONE* **9**, e84006 (2014).
- K. Yano, J. D. Stevens, L. J. V. Compagno, *J. Fish Biol.* **70**, 374–390 (2007).
- J. Nielsen, R. B. Hedeholm, M. Simon, J. F. Steffensen, *Polar Biol.* **37**, 37–46 (2014).
- B. C. McMeans, J. Svavarsson, S. Dennard, A. T. Fisk, *Can. J. Fish. Aquat. Sci.* **67**, 1428–1438 (2010).
- J. H. Hansen, R. B. Hedeholm, K. Sünksen, J. T. Christensen, P. Grønkjær, *Mar. Ecol. Prog. Ser.* **467**, 47–59 (2012).
- L. J. V. Compagno, Ed., *FAO Species Catalogue. Vol. 4: Sharks of the World. An Annotated and Illustrated Catalogue of the Shark Species Known to Date. Part 1. Haxanchiformes to Lamniformes* (FAO Fisheries Synopsis, Food and Agriculture Organization of the United Nations, ed. 4, 1984).
- M. P. Heide-Jørgensen, J. Teilman, *Biosci* **39**, 195–212 (1994).
- W. N. Joyce et al., *ICES J. Mar. Sci.* **59**, 1263–1269 (2002).
- S. E. Campana, C. Jones, G. A. McFarlane, S. Myklevoll, *Environ. Biol. Fishes* **77**, 327–336 (2006).
- J. A. Estrada, A. N. Rice, L. J. Natanson, G. B. Skomal, *Ecology* **87**, 829–834 (2006).
- R. E. A. Stewart, S. E. Campana, C. M. Jones, B. E. Stewart, *Can. J. Zool.* **84**, 1840–1852 (2006).
- P. J. Reimer et al., *Radiocarbon* **55**, 1869–1887 (2013).
- J. D. Scourse et al., *Radiocarbon* **54**, 165–186 (2012).
- C. Bronk Ramsey, *Radiocarbon* **37**, 425–430 (1995).
- C. Bronk Ramsey, *Quat. Sci. Rev.* **27**, 42–60 (2008).
- C. Bronk Ramsey, S. Lee, *Radiocarbon* **55**, 720–730 (2013).
- C. Bronk Ramsey, *Radiocarbon* **51**, 1023–1045 (2009).
- J. T. Christensen, K. Richardson, *Mar. Ecol. Prog. Ser.* **368**, 1–8 (2008).
- P. G. Butler et al., *Earth Planet. Sci. Lett.* **279**, 230–241 (2009).
- A. T. Fisk, S. A. Tittlemier, J. L. Pranschke, R. J. Norstrom, *Ecology* **83**, 2162–2172 (2002).
- P. G. Butler, A. D. Wanamaker Jr., J. D. Scourse, C. A. Richardson, D. J. Reynolds, *Palaeogeogr. Palaeoclimatol.* **373**, 141–151 (2013).
- R. B. Stouby, *Eksportkrone for Skidtfisk* (Eksportrådet: The Trade Council, Danish Ministry of Foreign Affairs 2, Copenhagen, Denmark, 2011).
- S. E. Campana, *Mar. Ecol. Prog. Ser.* **150**, 49–56 (1997).
- J. M. Kalish, R. Nydal, K. H. Nedreaas, G. S. Burr, G. L. Eine, *Radiocarbon* **43**, 843–855 (2001).
- M. A. Treble, S. E. Campana, R. J. Wastle, C. N. Jones, J. Boje, *Can. J. Sci. Aquat. Sci.* **65**, 1047–1059 (2008).

ACKNOWLEDGMENTS

We are grateful for the contributions from M. B. Backe throughout the manuscript. We thank the Commission of Scientific Investigations in Greenland (KVUG), Save Our Seas Foundation, National Geographic Foundation, Carlsberg Foundation, Danish Centre for Marine Research, Den Blå Planet–National Aquarium of Denmark, Greenland Institute of Natural Resources (GINR), and the Danish Council for Independent Research for financial support. We thank GINR, the University of Copenhagen and the TUNU Programme (UIT, The Arctic University of Norway) for ship time. We are grateful for the collaboration with K.P. Lange. We thank the crews of the RV *Pamiut*, RV *Dana*, RV *Helmer Hansen*, RV *Sanna*, and RV *Porsild*. Three anonymous reviewers provided helpful comments and discussion that improved earlier versions of the manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/702/suppl/DC1
Material and Methods
Supplementary Text
Figs. S1 and S2
Tables S1 and S2
References (39–50)

29 December 2015; accepted 10 June 2016
10.1126/science.aaf1703

HYDROLOGY

Continental patterns of submarine groundwater discharge reveal coastal vulnerabilities

Audrey H. Sawyer,^{1*} Cédric H. David,² James S. Famiglietti²

Submarine groundwater discharge (SGD) delivers water and dissolved chemicals from continents to oceans, and its spatial distribution affects coastal water quality. Unlike rivers, SGD is broadly distributed and relatively difficult to measure, especially at continental scales. We present spatially resolved estimates of fresh (land-derived) SGD for the contiguous United States, based on historical climate records and high-resolution hydrographic data. Climate controls regional patterns in fresh SGD, while coastal drainage geometry imparts strong local variability. Because the recharge zones that contribute fresh SGD are densely populated, the quality and quantity of fresh SGD are both vulnerable to anthropogenic disturbance. Our analysis unveils hot spots for contaminant discharge to marine waters and saltwater intrusion into coastal aquifers.

Submarine groundwater discharge (SGD) influences global geochemical cycles and coastal water quality by delivering chemical compounds and dissolved ions from land to sea (1, 2). SGD includes two primary components: fresh, land-derived groundwater that infiltrates on land, and salty, ocean-derived groundwater that infiltrates offshore and returns to the sea (3). Although small in volume, fresh SGD exports naturally derived elements such as calcium and silicate at rates that rival those of rivers (4, 5). Fresh SGD is sensitive to human disturbance, and mixtures of fresh and saline SGD transport nutrients and other contaminants offshore. Therefore, the spatial distribution of fresh SGD has a direct impact on patterns of coastal water quality. High rates of nutrient-rich SGD, for example, contribute to harmful algal blooms and hypoxia (6, 7). SGD patterns also influence ocean temperature (8) and alkalinity (9), which are key controls on marine ecological and biogeochemical processes.

To understand the influence of fresh SGD on biogeochemical cycles and coastal water quality, rate assessments are needed at global and local scales. However, fresh SGD is difficult and costly to measure, and observations are scarce (8). Fresh SGD is heterogeneous and diffuse, unlike river discharge, which is concentrated at discrete and readily measurable points along the coast. Although many techniques exist to measure fresh SGD, most measurements are focused on a handful of well-studied, easily accessible locations. The majority of these locations are on the Atlantic coast of the United States (8) (Fig. 1). In the absence of measurements, water budgets have been used to map fresh SGD at high resolution over small

coastal regions (10) and low resolution across the global oceans (11). These disparate scales of analysis leave critical gaps in our understanding of SGD. High-resolution estimates are needed across large regions to reveal relationships among climate, geology, and SGD and to identify where coastal water resources are vulnerable to degradation.

Here, we present high-resolution continental-scale estimates of fresh SGD across the contiguous United States. Our estimates are based on a simple water budget analysis and state-of-the-art continental-scale hydrography and climate data sets. Recharge zones, or contributing areas, for fresh SGD are defined with high-resolution hydrographic data. We assume that recharge zones are the wedge-shaped land areas outside stream catchments where water flows directly to the coast (Fig. 2, inset). We then assume that recharge across these coastal catchments is the component of precipitation that infiltrates and would become base flow to a stream, if a stream were present, but instead flows to the coast. This recharge rate is derived from three decades of reanalysis of climatic data (12).

To validate our estimates, we compiled 18 local estimates of SGD across the United States (12). We sought representation from diverse locations along the Pacific, Gulf, and Atlantic coasts and favored studies that used direct near-shore measurements to estimate the fresh component of SGD wherever possible. We excluded sea-based measurements using radon or radium tracer techniques, which can capture a large saline component of SGD (3, 13, 14). Sea-based measurements sometimes predict substantially larger SGD rates than near-shore measurements, simulations, and water budget-based estimates (14, 15). Our predicted SGD rates are correlated with local estimates but are consistently lower (Fig. 1). The magnitude of discrepancy shows no apparent relationship with geology, climate, land use, or population

density. Some field measurements may overestimate fresh SGD, because sites are often selected where fresh SGD is likely to be focused (for example, in permeable sands or bay heads where groundwater flow paths converge). Although field measurements may overestimate fresh SGD, our approach likely underestimates them because coastal recharge zones could import groundwater from upland catchments (16). These potential additional groundwater sources are not included in our analysis. Furthermore, fresh SGD estimates from water balance approaches tend to decrease with increasing spatial resolution (17), and our analysis uses high-resolution hydrography data. The approach nevertheless allows for unprecedented mapping of fresh SGD.

At the local scale, our analysis exposes a strong heterogeneity in fresh SGD rates (Fig. 2, expanded map). This heterogeneity can be explained by the variability in land area that contributes groundwater to a given length of coastline. We define the coastal drainage length, which represents the average distance that groundwater travels from its point of recharge to the coast (Fig. 2, inset). The drainage length equals the recharge area for fresh SGD divided by the length of coastline where discharge occurs. It varies strongly with local topography and locations of coastal rivers (Fig. 2, expanded map). As a result, the spatial variability in fresh SGD over a typical 100-km segment of shoreline is almost as large as the variability at the continental scale. For example, the coefficient of variation for SGD within 100 km of San Francisco is 0.76, whereas its coefficient of variation along the entire West Coast is 1.10 (Fig. 2). Because of this strong local variability, SGD measurements at a single site cannot be extrapolated to other nearby sites with high confidence. Moreover,

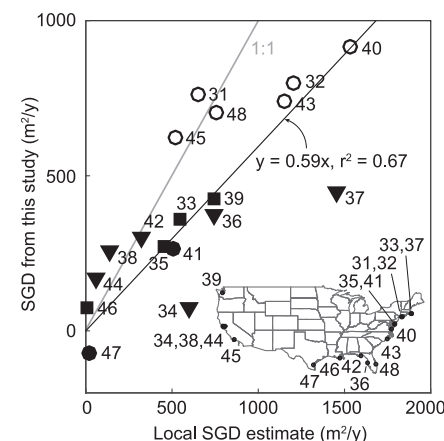


Fig. 1. SGD estimates from this study and the literature. Our SGD values correlate well with local estimates (12) but are generally lower. Circles are from seepage meters, squares are from water budget analysis, and triangles are from multiple methods. Solid symbols indicate fresh SGD estimates. Open circles indicate total (fresh and saline) SGD. We calculated total SGD rates from our fresh rates using the relation $\text{SGD}_{\text{total}} = 1.1 \text{SGD}_{\text{fresh}} + 470 \text{ m}^3/\text{year}$ (14, 30).

¹School of Earth Sciences, Ohio State University, Columbus, OH 43210, USA.

²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA.

*Corresponding author. Email: sawyer.143@osu.edu

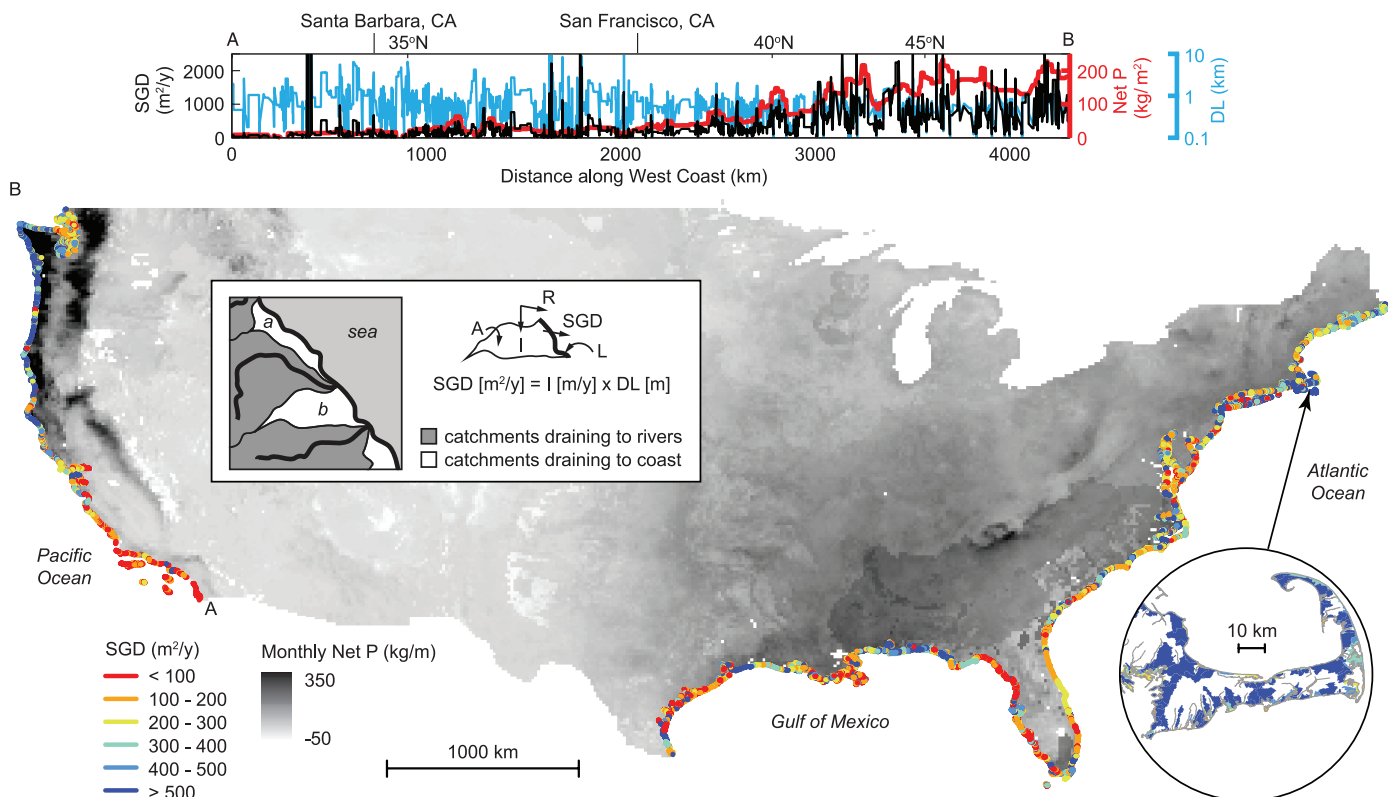


Fig. 2. Map of fresh SGD rates along the contiguous United States coast. On the West Coast, fresh SGD increases from south to north (point A to point B) with net precipitation (Net P) while drainage length (DL) remains consistent. The shape of recharge zones (inset in map) dictates local variability; recharge zone a has shorter DL, whereas zone b has longer DL. Fresh SGD is calculated from the product of infiltrating precipitation (I) and DL ($DL=A/L$, where A is recharge zone area and L is coastal length). Noninfiltrating runoff (R) does not contribute to fresh SGD. Expanded view of Cape Cod, Massachusetts shows coastal recharge zones colored by rate of fresh SGD.

human modifications to coastal drainage networks affect patterns of fresh SGD. For example, fresh SGD rates are low in some areas of Florida with highly altered drainage networks (fig. S2).

At the continental scale, patterns in fresh SGD depend on both drainage geometry and climate. The influence of climate is clear along the West Coast from Southern California to Washington (Fig. 2), where net precipitation and fresh SGD both increase by more than 90%, but coastal drainage length is consistent. Meanwhile, the influence of drainage length is evident across East and West coasts. For example, net precipitation is similar in the Pacific Northwest and the mid-Atlantic, but fresh SGD rates are ~50% greater in the Pacific Northwest because of the abundance of long coastal drainage lengths in steep terrain (Fig. 2 and fig. S1).

At the continental scale, recharge areas for fresh SGD constitute a small portion of the total land area (0.4% of the contiguous United States). However, these areas drain more water than the continental interior on an areal basis because they receive more net precipitation. The total volumetric rate of fresh SGD from the contiguous United States to the oceans is 15 ± 4 cubic km/year, or <1% of total land runoff (18). Because this simple water budget analysis may underestimate fresh

SGD by up to 40% (Fig. 1), rates may be as high as 25 ± 7 cubic km/year, or <2% of runoff. These continental-scale estimates are in line with previous estimates for the contiguous United States [1 to 10% of runoff (19)] and the world [6% of runoff (11)]. Note that saline SGD is substantially greater than the fresh component estimated here (3) and may be as large as 300 to 400% of global runoff (20), but the fresh component is most vulnerable to contamination and other anthropogenic disturbances. Although volumetrically small, fresh SGD can carry large contaminant mass loads. For example, in some parts of the world, fresh SGD delivers up to 30 times as much nitrogen to the coast as rivers (21, 22).

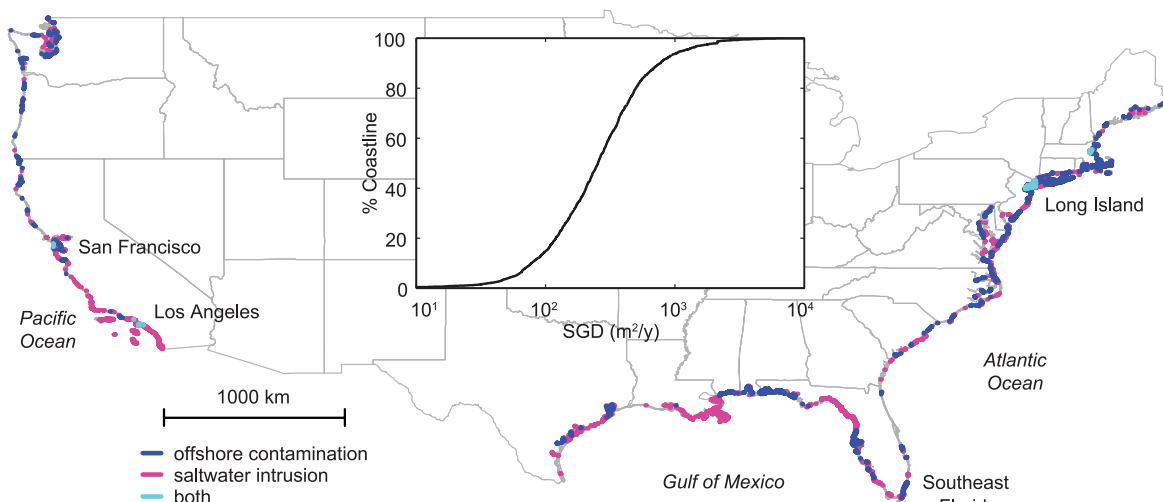
The average volumetric flux of fresh SGD per unit length of coastline is 420 m³/year, but rates span orders of magnitude (Fig. 3, inset). Although SGD is ubiquitous, concentrated discharge zones contribute the majority of fresh SGD to the oceans: Half of all fresh SGD is focused along only 14% of the coast. Interestingly, rates of fresh SGD follow a log-normal distribution (Fig. 3, inset), like permeability values for the shallow earth (23). Permeability is difficult to measure because it ranges by orders of magnitude and is scale-dependent. A particular strength of our analysis is that it does not require permeability data but relies

instead on standardized topographic and climatic data sets.

Most of the global population lives near and depends on coastal water resources and fisheries. Thus, high-resolution data sets are imperative for identifying coastal waters that may be vulnerable to “hidden” contaminant loads from fresh SGD. Within the contiguous United States, 3% of the population inhabits recharge areas for fresh SGD, which represent only 0.4% of the total land area. Although 72% of recharge areas were undeveloped as of 2011, conversion to agricultural and urban land use is ongoing. With coastal land development, nutrient loads to groundwater from septic tanks and fertilizers are increasing. Making matters worse, wetland loss reduces coastal resilience to contaminant loading because wetlands are efficient contaminant filters. Regions with above-average fresh SGD and land use development are particularly vulnerable to groundwater-borne contamination, and these regions represent 12% of the coastline. Vulnerable regions include the northern Gulf Coast from Mississippi to the Florida Panhandle, northern Atlantic Coast, and Pacific Northwest (Fig. 3). These regions have previously been shown to have high potential nitrogen inputs from SGD to coastal waters (24). Vulnerable regions should

Fig. 3. Coastal vulnerability map.

Vulnerability to offshore contamination (dark blue) is identified where the rate of fresh SGD is above average ($420 \text{ m}^2/\text{year}$) and developed or agricultural land use is above average (27.7%). Vulnerability to saltwater intrusion (magenta) is identified where low fresh SGD or high groundwater extraction may cause complete saltwater invasion. Light blue areas of coastlines are vulnerable to both offshore contamination and saltwater intrusion. Inset shows histogram of fresh SGD rates for the contiguous United States.



be monitored for direct nutrient inputs from fresh SGD.

High-resolution maps of SGD are also useful for assessing vulnerability of coastal aquifers to saltwater intrusion (Fig. 3). In populated areas, groundwater extraction subtracts from the recharge available for fresh SGD and can lead to saltwater intrusion. Most coastal aquifers are more sensitive to groundwater extraction than sea level rise (25). We predicted vulnerability to saltwater intrusion using an analytical solution for the position of the freshwater-saltwater interface (26, 27), assuming a population-dependent groundwater extraction rate that directly subtracts from the rate of fresh SGD (12). Coastlines are considered vulnerable where the toe of the freshwater-saltwater interface reaches the groundwater divide of the coastal aquifer, which implies imminent and full saltwater intrusion. These regions represent 9% of the coastline and include confirmed locations of saltwater intrusion such as Long Island, New York (28) and Los Angeles, California (29) (Fig. 3).

Because of the highly heterogeneous nature of fresh SGD, neighboring coastal zones can be vulnerable to saltwater intrusion and discharge of groundwater-borne contaminants to the ocean. Regions of mixed vulnerability include southeastern Florida and New Jersey, among many others. Only a small fraction of coastline (<1%) is dually vulnerable to both saltwater intrusion and offshore contamination, including the heavily developed and populated areas of San Francisco, Los Angeles, and Long Island (Fig. 3). If these urban areas rely primarily on groundwater to meet their resource demands, the resulting deductions to fresh SGD may cause full saltwater intrusion. Conversely, effective groundwater management may sustain fresh SGD rates, but contaminant loads to the coast may be high.

As the resolution of global hydrographic data improves, this same approach can be used to pre-

dict global distributions of fresh SGD and vulnerabilities in coastal water quality. Vulnerable regions will shift and likely grow with coastal land use change, population growth, climate change, and sea level rise. In many areas, rates of fresh SGD will decrease as impervious pavement expands and groundwater withdrawals increase. Regions with high rates of fresh SGD that are currently vulnerable to offshore contamination may instead become vulnerable to saltwater intrusion. Because rates of fresh SGD are highly heterogeneous, spatial estimates are imperative for identifying monitoring needs and assessing threats to coastal water quality on both sides of the land-sea boundary, in onshore aquifers and marine surface waters.

REFERENCES AND NOTES

- W. S. Moore, *Nature* **380**, 612–614 (1996).
- W. S. Moore, *Annu. Rev. Mar. Sci.* **2**, 59–88 (2010).
- W. C. Burnett, H. Bokuniewicz, M. Huettel, W. S. Moore, M. Taniguchi, *Biogeochemistry* **66**, 3–33 (2003).
- S. D. Rad, C. J. Allegre, P. Louvat, *Earth Planet. Sci. Lett.* **262**, 109–124 (2007).
- H. H. Schopka, L. A. Derry, *Earth Planet. Sci. Lett.* **339–340**, 67–78 (2012).
- C. M. Hu, F. E. Muller-Karger, P. W. Swarzenski, *Geophys. Res. Lett.* **33**, L11601 (2006).
- H. W. Paerl, *Limnol. Oceanogr.* **42**, 1154–1165 (1997).
- W. C. Burnett *et al.*, *Sci. Total Environ.* **367**, 498–543 (2006).
- T. Cyronak, I. R. Santos, D. V. Erler, B. D. Eyre, *Biogeosciences* **10**, 2467–2480 (2013).
- G. Destouni, F. Hannerz, C. Prieto, J. Jarsjo, Y. Shibuo, *Global Biogeochem. Cycles* **22**, GB4003 (2008).
- I. S. Zektser, H. A. Loaiciga, *J. Hydrol.* **144**, 405–427 (1993).
- See supplementary materials on Science Online.
- H. A. Michael, M. A. Charette, C. F. Harvey, *Mar. Chem.* **127**, 100–114 (2011).
- C. Prieto, G. Destouni, *Geophys. Res. Lett.* **38**, L01402 (2011).
- P. L. Younger, *Nature* **382**, 121–122 (1996).
- M. F. Schaller, Y. Fan, *J. Geophys. Res.* **114**, D04103 (2009).
- G. Destouni, Y. Shibuo, J. Jarsjo, *Geophys. Res. Lett.* **35**, L18401 (2008).
- W. A. Gebert, D. J. Graczyk, W. R. Krug, "Average Annual Runoff in the United States, 1951–80" (U.S. Geological Survey Hydrologic Atlas, 1987).
- I. S. Zektser, V. A. Ivanov, A. V. Meskheteli, *J. Hydrol.* **20**, 1–36 (1973).
- E. Y. Kwon *et al.*, *Geophys. Res. Lett.* **41**, 8438–8444 (2014).
- C. P. Slomp, P. Van Cappellen, *J. Hydrol.* **295**, 64–86 (2004).
- I. Valiela *et al.*, *Biogeochemistry* **10**, 177–197 (1990).
- T. Gleeson *et al.*, *Geophys. Res. Lett.* **38**, L02401 (2011).
- A. H. W. Beusen, C. P. Slomp, A. F. Bouwman, *Environ. Res. Lett.* **8**, 034035 (2013).
- G. Ferguson, T. Gleeson, *Nat. Clim. Change* **2**, 342–345 (2012).
- K. Mazi, A. D. Koussis, G. Destouni, *Environ. Res. Lett.* **8**, 014001 (2013).
- K. Mazi, A. D. Koussis, G. Destouni, *Hydrol. Earth Syst. Sci.* **18**, 1663–1677 (2014).
- P. M. Barlow, *Ground Water in Freshwater-Saltwater Environments of the Atlantic Coast* (U.S. Geological Survey, 2003).
- P. M. Barlow, E. G. Reichard, *Hydrogeol. J.* **18**, 247–260 (2010).
- G. Destouni, C. Prieto, *Biogeochemistry* **66**, 171–186 (2003).

ACKNOWLEDGMENTS

We thank three anonymous reviewers for their suggestions, and M. Durand, H. Michael, C. Russoniello, and J. Heiss for discussions. Supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA, and grants from the NASA SWOT and Sea Level Science Teams (C.H.D. and J.S.F.); NSF grant EAR-1446724; and the Ohio State University School of Earth Sciences. The authors declare no competing interests. Fresh SGD rates and associated data are freely available at <http://dx.doi.org/10.5281/zenodo.58871>.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/705/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 and S2
References (31–53)

9 May 2016; accepted 18 July 2016
Published online 4 August 2016
10.1126/science.aag1058

NEURODEGENERATION

Spt4 selectively regulates the expression of *C9orf72* sense and antisense mutant transcripts

Nicholas J. Kramer,^{1,2*} Yari Carlomagno,^{3*} Yong-Jie Zhang,^{3*} Sandra Almeida,⁴ Casey N. Cook,³ Tania F. Gendron,³ Mercedes Prudencio,³ Marka Van Blitterswijk,³ Veronique Belzil,³ Julien Couthouis,¹ Joseph West Paul III,¹ Lindsey D. Goodman,⁵ Lillian Daugherty,³ Jeannie Chew,³ Aliesha Garrett,³ Luc Pregent,³ Karen Jansen-West,³ Lilia J. Tabassian,³ Rosa Rademakers,³ Kevin Boylan,⁶ Neill R. Graff-Radford,⁶ Keith A. Josephs,⁷ Joseph E. Parisi,⁷ David S. Knopman,⁷ Ronald C. Petersen,⁷ Bradley F. Boeve,⁷ Ning Deng,¹ Yanan Feng,¹ Tzu-Hao Cheng,⁸ Dennis W. Dickson,³ Stanley N. Cohen,¹ Nancy M. Bonini,⁵ Christopher D. Link,⁹ Fen-Biao Gao,⁴ Leonard Petrucelli,^{3†} Aaron D. Gitler^{1†}

An expanded hexanucleotide repeat in *C9orf72* causes amyotrophic lateral sclerosis and frontotemporal dementia (c9FTD/ALS). Therapeutics are being developed to target RNAs containing the expanded repeat sequence (GGGGCC); however, this approach is complicated by the presence of antisense strand transcription of expanded GGCCCC repeats. We found that targeting the transcription elongation factor Spt4 selectively decreased production of both sense and antisense expanded transcripts, as well as their translated dipeptide repeat (DPR) products, and also mitigated degeneration in animal models. Knockdown of SUPT4H1, the human Spt4 ortholog, similarly decreased production of sense and antisense RNA foci, as well as DPR proteins, in patient cells. Therapeutic targeting of a single factor to eliminate c9FTD/ALS pathological features offers advantages over approaches that require targeting sense and antisense repeats separately.

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are devastating neurodegenerative diseases. The *C9orf72* mutation (1, 2) is a common cause of both ALS and FTD (c9FTD/ALS), and intense efforts are exploring disease mechanisms and developing therapeutic strategies (3, 4). The *C9orf72* gene harbors a hexanucleotide repeat, GGGGCC, located in the first intron. In c9FTD/ALS patients, this hexanucleotide repeat tract is expanded to hundreds or even thousands of repeats. Several hypotheses have emerged to explain how *C9orf72* mutations cause neurodegeneration. RNAs transcribed from the GGGGCC repeats that accumulate as foci in the nucleus and cytoplasm could cause disease by sequestering RNA-binding proteins and splicing factors via an RNA

toxicity mechanism (5–7). The transcribed nucleotide repeats are also substrates for an unconventional form of translation, termed repeat-associated non-ATG (RAN) translation (8–10). The resulting dipeptide repeat (DPR) proteins could cause disease through proteotoxic mechanisms (11). Lastly, the expanded repeats reduce *C9orf72* mRNA expression (12, 13), which may cause a loss of *C9orf72* function. A therapeutic approach that uses antisense oligonucleotides (ASOs) to target GGGGCC repeat-containing RNA transcripts for degradation is being pursued (5, 14, 15). However, RNA foci containing GGCCCC repeats transcribed in the antisense direction also accumulate abundantly in c9FTD/ALS (14, 16). Thus, it will be important to consider strategies that target both sense and antisense repeat-containing RNAs.

Spt4 is a highly conserved transcription elongation factor that, together with its binding partner Spt5, regulates RNA polymerase II processivity (17–20). A recent genetic screen in yeast revealed that mutation of Spt4 selectively reduced the transcription of long CAG trinucleotide repeats associated with Huntington's disease (21). Inhibiting the mammalian ortholog of Spt4, SUPT4H1, also decreased production of expanded CAG-derived polyQ aggregates in murine striatal neurons. Notably, lowering levels of Spt4 did not affect expression of proteins containing short CAG repeat regions. Moreover, RNA-sequencing (RNA-seq) analysis found only a limited effect of *SPT4* deletion on normal gene expression throughout

the transcriptome (21). Thus, Spt4 acts as a specialized transcription factor selectively required for the expression of long trinucleotide repeat-containing transcripts.

Because Spt4 inhibition lowers the expression of disease-associated trinucleotide repeat expansions (21), we hypothesized that Spt4 inhibition might reduce expression of expanded sense and antisense hexanucleotide repeats in mutant *C9orf72*. If so, it could overcome the current hurdles of designing and optimizing two strategies to target sense and antisense repeats separately. To address this hypothesis, we began by testing the effect of *SPT4* deletion in yeast (*Saccharomyces cerevisiae*) expressing *C9orf72* hexanucleotide repeats. We transformed wild-type (WT) yeast cells with a galactose-inducible expression construct harboring either short (2) or expanded (66) sense GGGGCC repeats or antisense GGCCCC repeats lacking an ATG start codon. Although yeast expressing expanded sense or antisense repeats formed cytoplasmic and perinuclear RNA foci, neither the short nor the expanded repeat constructs were toxic, as assessed by serial dilution growth analysis of yeast strains (fig. S1, A to D). Using quantitative reverse transcription-polymerase chain reaction (qRT-PCR), we next evaluated the effect of *SPT4* deletion (*spt4Δ*) on transcription of sense and antisense repeats. Whereas *spt4Δ* had no effect on short, nonexpanded repeats, we observed a significant decrease in expression of (GGGGCC)₆₆ or (GGCCCC)₆₆ transcripts in *spt4Δ* cells relative to WT (Fig. 1A). Moreover, *SPT4* deletion completely blocked RNA foci formation in yeast expressing (GGGGCC)₆₆ or (GGCCCC)₆₆ (Fig. 1, B to D). Expressing Spt4 from a plasmid in *spt4Δ* cells restored sense and antisense RNA foci formation (Fig. 1, B to D), confirming the requirement of Spt4 for the production of transcripts containing expanded *C9orf72* repeats and the subsequent accumulation of these RNA transcripts into foci.

In addition to RNA foci, another pathological feature of c9FTD/ALS is the presence of DPR proteins produced by RAN translation (8, 9, 22). To investigate whether RAN translation occurs in yeast expressing expanded GGGGCC repeats, we used an immunoassay to detect expression of a DPR protein produced from the glycine-proline (GP) reading frame. Poly(GP) was not detected in lysates of yeast cells expressing 2 or 66 GGGGCC repeats from a low-copy (CEN) expression plasmid; however, robust poly(GP) expression was observed in lysates of yeast cells in which 40 GGGGCC repeats were expressed at higher levels from a high-copy (2-μm) expression plasmid (Fig. 1E and fig. S1C). These data are consistent with the prior finding that the abundance of poly(GP) proteins correlates with levels of repeat-containing transcripts (11, 12). Notably, deleting *SPT4* in (GGGGCC)₄₀ expressing yeast significantly decreased levels of RAN-translated poly(GP) (Fig. 1E). This effect was specific to the expanded nucleotide repeat construct (as in Fig. 1A), because *SPT4* deletion did not suppress toxicity of three non-repeat-containing disease-associated proteins [TDP-43, FUS/TLS, and poly(PR)] (fig. S2A). Further, *SPT4*

¹Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. ²Neurosciences Graduate Program, Stanford University School of Medicine, Stanford, CA, USA. ³Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA. ⁴Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA. ⁵Department of Biology and Neuroscience Graduate Program, University of Pennsylvania, Philadelphia, PA, USA. ⁶Department of Neurology, Mayo Clinic, Jacksonville, FL, USA. ⁷Department of Neurology, Mayo Clinic, Rochester, MN, USA. ⁸Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan, Republic of China. ⁹Department of Integrative Physiology, University of Colorado, Boulder, CO, USA.

*These authors contributed equally to this work. †Corresponding author. Email: agitler@stanford.edu (A.D.G.); petrucelli.leonard@mayo.edu (L.P.)

deletion did not affect expression of green fluorescent protein under the control of the same galactose-inducible promoter or expression of the endogenous yeast glyceraldehyde-3-phosphate dehydrogenase (GAPDH) protein (fig. S2B). Thus, two key components of c9FTD/ALS pathology (RNA foci and DPR proteins) are recapitulated in yeast models and are mitigated by Spt4 depletion.

We next tested the effect of modulating Spt4 levels in vivo using animal models. Using the synaptobrevin (*snb-1*) promoter to drive neuronal expression of (GGGGCC)₆₆ (fig. S3A), we developed a nematode (*Caenorhabditis elegans*) model of the *C9orf72* repeat expansion characterized by RNA foci formation (Fig. 2A and fig. S3, B and C), DPR protein production (Fig. 2C), and a significant reduction in life span (Fig. 2D and fig. S3G). To determine the effect of altered Spt4/SUPT4H1 expression levels, we crossed (GGGGCC)₆₆-expressing worms to worms expressing human SUPT4H1 also driven by the *snb-1* promoter (Fig. 2, A to D, and fig. S3, D to F) or fed RNAi to knock down endogenous Spt4 (Fig. 2, E to H). We observed an increase in GGGGCC repeat RNA and poly(GP) expression, as well as enhanced toxicity in (GGGGCC)₆₆ worms in

the presence of exogenous SUPT4H1 (Fig. 2, B to D, and fig. S3, E to G). Conversely, reduction in endogenous Spt4 levels decreased both GGGGCC repeat RNA and poly(GP) levels and also improved the survival of (GGGGCC)₆₆ worms (Fig. 2, E to H, and fig. S3G). Similar modulation of toxicity was observed in transgenic *Drosophila* engineered to express 6, 29, or 49 pure GGGGCC repeats under the control of the *gmr-GAL4* driver. Consistent with previous reports (22–26), expression of (GGGGCC)₄₉ caused severe eye degeneration (Fig. 2I and fig. S4), whereas the shorter expanded repeat, (GGGGCC)₂₉, caused a mild phenotype with a modest disruption of the external eye and significant thinning of the retinal ganglia. The unexpanded repeat (GGGGCC)₆ caused no observable effect in the retina. *Spt4* RNAi partially suppressed the degenerative phenotype of the external and internal eye in (GGGGCC)₄₉-expressing flies and almost completely suppressed the retinal thinning normally observed in (GGGGCC)₂₉-expressing flies (Fig. 2J and fig. S4, B to D). As in yeast, the effect of Spt4 depletion was specific to expression of expanded GGGGCC repeats because the expression of a

control β -galactosidase transgene was unaffected (fig. S4E). Furthermore, reducing the expression of Spt4 significantly increased the life span of adult flies ubiquitously expressing (GGGGCC)₄₉ repeats (Fig. 2K). Thus, the level of Spt4 expression influences the pathogenicity of *C9orf72* GGGGCC repeats in the nervous systems of both *C. elegans* and *Drosophila*.

The experiments in yeast, nematodes, and flies are consistent with a role for Spt4 in regulating expression of *C9orf72* hexanucleotide repeat expansions. These experiments all relied on expression of GGGGCC or GGCCCC repeats from plasmids or transgenes. To extend our findings, we tested the ability of SUPT4H1 to regulate expression of endogenous *C9orf72* repeats in cells obtained from humans afflicted with ALS. We first analyzed cultured fibroblast cells derived from healthy subjects or from ALS patients with or without *C9orf72* repeat expansions (fig. S5A). Foci of sense and antisense repeat-containing RNA (fig. S5B) and poly(GP) proteins (fig. S5C) were selectively detected in c9ALS fibroblasts. Consistent with previous findings in brain tissues from c9FTD/ALS patients (11), the abundance of poly(GP) proteins

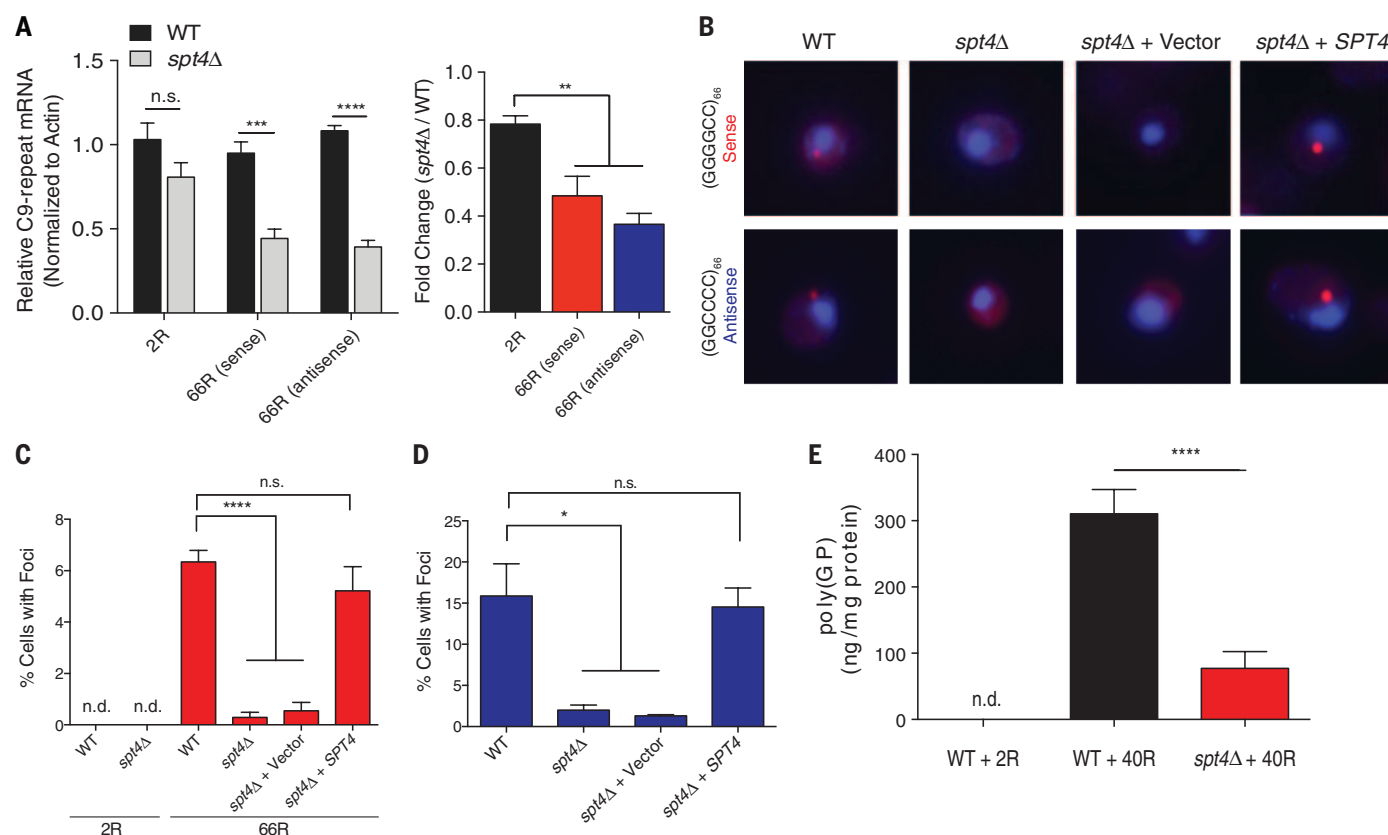


Fig. 1. Spt4 is required for *C9orf72* mutant strand transcription, RNA foci formation, and RAN translation in *S. cerevisiae*. Yeast were transformed with plasmids expressing sense (GGGGCC)_n or antisense (GGCCCC)_n *C9orf72* repeats of varying lengths (2R or 66R) but lacking ATG translation start codons under the control of a galactose-inducible promoter. (A) qRT-PCR analysis of C9-repeat RNA levels in WT and *spt4Δ* yeast after 6-hour galactose inductions. C9-repeat RNA levels were normalized to levels of yeast actin, then quantified relative to WT yeast (two-tailed unpaired *t* tests, *****P* < 0.0001, ****P* < 0.001; n.s., not significant). Data expressed as relative fold changes,

spt4Δ/WT [one-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test, ***P* < 0.01]. (B) Sense and antisense RNA foci were detected by fluorescence in situ hybridization (FISH). Percentages of yeast cells containing sense (C) and antisense (D) foci were quantified across genotypes [blue, 4',6-diamidino-2-phenylindole (DAPI); red, C9-repeat LNA probe; one-way ANOVA with Dunnett's multiple comparisons test, *****P* < 0.0001, **P* < 0.05; n.s., not significant; n.d., not determined]. (E) Sandwich immunoassay detection of RAN-translated poly(GP) in WT and *spt4Δ* yeast (two-tailed unpaired *t* test *****P* < 0.0001).

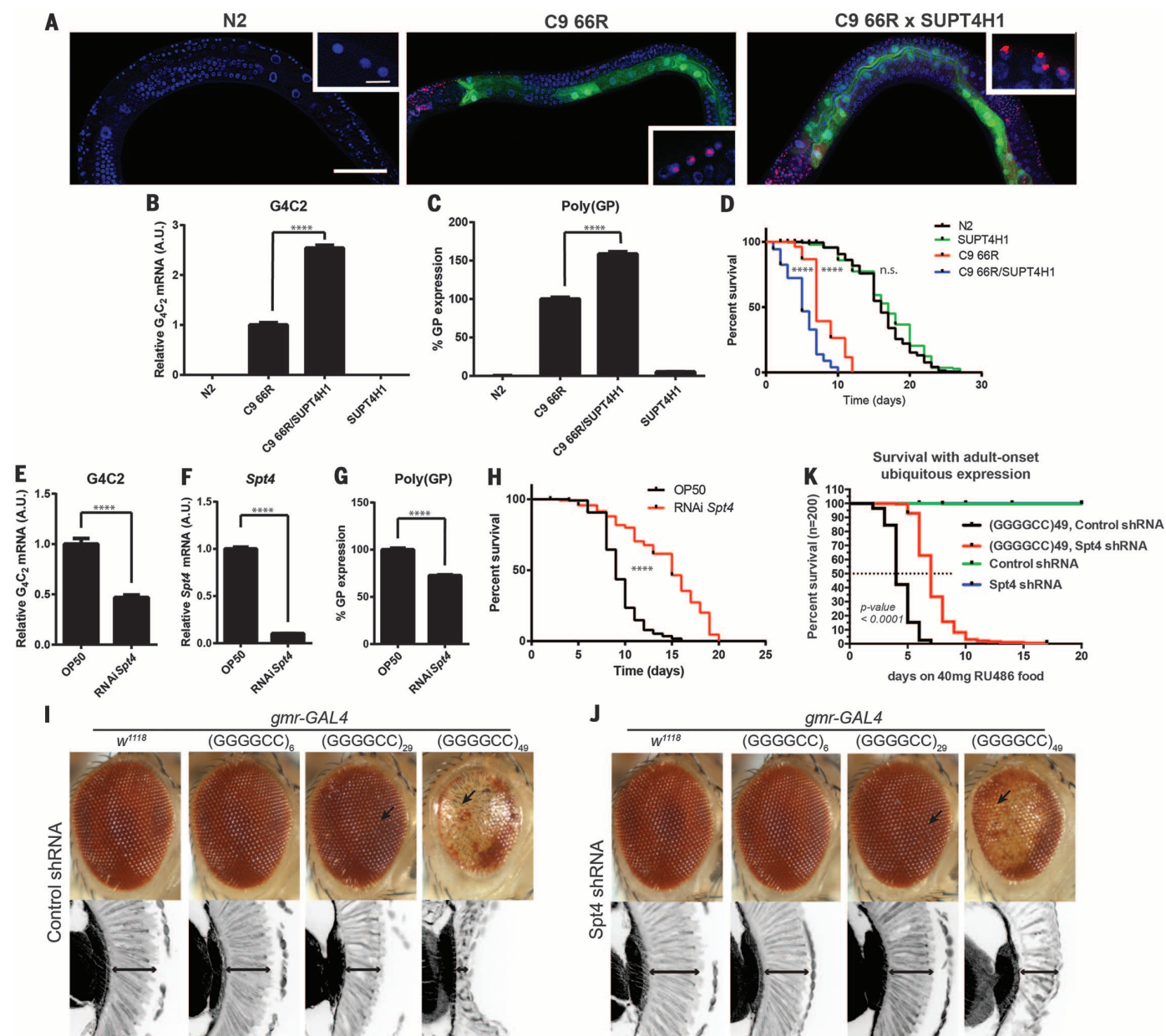


Fig. 2. Spt4 modulation influences pathogenicity of GGGGCC repeats in *C. elegans* and mitigates neurodegeneration in a *Drosophila* model of c9FTD/ALS. (A) RNA FISH reveals GGGGCC RNA foci formation (red) in normal (N2) worms, (GGGGCC)₆₆-expressing worms (C9 66R), and worms expressing both (GGGGCC)₆₆ and human SUPT4H1 (C9 66R × SUPT4H1). The intestinal marker (green) denotes the presence of the (GGGGCC)₆₆ transgene, and DAPI (blue) indicates labeled nuclei (scale bar: low magnification 50 μ m, high magnification 5 μ m). (B and C) Expression of GGGGCC RNA levels quantified by qPCR relative to levels in (GGGGCC)₆₆-expressing worms (two-tailed unpaired *t* test, *****P* < 0.0001) (B) and poly(GP) levels quantified by immunoassay (two-tailed unpaired *t* test, *****P* < 0.0001) (C). (D) A life-span assay assessed the impact of (GGGGCC)₆₆ and/or SUPT4H1 overexpression on survival (N2 *n* = 145 worms, C9 66R *n* = 160, C9 66R × SUPT4H1 *n* = 122, SUPT4H1 *n* = 144; log-rank (Mantel-Cox) test *****P* < 0.0001; n.s., not significant. See also fig. S3G). (E to H) qPCR measured GGGGCC (E) and Spt4 mRNA (F) levels in (GGGGCC)₆₆-expressing worms fed control (OP50) or RNAi targeting Spt4 (two-tailed unpaired *t* test, *****P* < 0.0001). (G) An immunoassay revealed the percentage change in poly(GP) levels after RNAi-mediated reduction in Spt4 expression (two-tailed unpaired *t* test, *****P* < 0.0001). (H) Spt4 RNAi improved survival in

(GGGGCC)₆₆ worms (OP50 *n* = 119 worms, RNAi Spt4 *n* = 120; log-rank (Mantel-Cox) test, *****P* < 0.0001. See also fig. S3G). (I and J) External and internal retinal structures of normal flies or flies expressing (GGGGCC)_n transgenes, with co-expression of (I) a control (luciferase) shRNA or (J) Spt4 shRNA. A *gmr-GAL4* eye-specific driver was used to express UAS-transgenes. (I) Control animals with driver alone (*w¹¹¹⁸*) or a short GGGGCC repeat (GGGGCC)₆ have normal eye structure. Animals expressing GGGGCC₂₉ show a mild disruption to the highly precise external ommatidial structure (arrows), and thinning and gaps in the internal retina, whereas animals expressing GGGGCC₄₉ have severely degenerate tissue both externally and internally. (J) Reduction in Spt4 has no effect on normal or GGGGCC₆-expressing animals and resulted in reduced toxicity in both the external and internal retinal structures of flies bearing 29 or 49 GGGGCC repeats. (K) Reducing the expression of Spt4 in (GGGGCC)₄₉ animals increased the life expectancy for animals expressing transgenes using a ubiquitous, drug-inducible driver, *da-GS*. Expression was induced in adult animals to avoid effects of expressing transgenes during development. Expression of the control (luciferase) shRNA or Spt4 shRNA had no effect on the life span over this time frame when expressed alone (curves superimposed). Statistical analysis was done using a log-ranked test comparing (GGGGCC)₄₉ animals expressing the control shRNA versus the Spt4 shRNA.

correlated with mRNA levels of *C9orf72* variant 3 (fig. S5G) but not of variant 1, variant 2, or total *C9orf72* mRNA (fig. S5, D to F) or with repeat length (fig. S5H).

Having established the occurrence of RAN translation and RNA foci formation in c9ALS fibroblasts, we evaluated whether depletion of SUPT4H1 and/or its binding partner, SUPT5H,

mitigates these pathological hallmarks. Treating cultured fibroblasts from three c9ALS patients with small interfering RNAs (siRNAs) against *SUPT4H1* and *SUPT5H* (siSUPT4H1 and siSUPT5H,

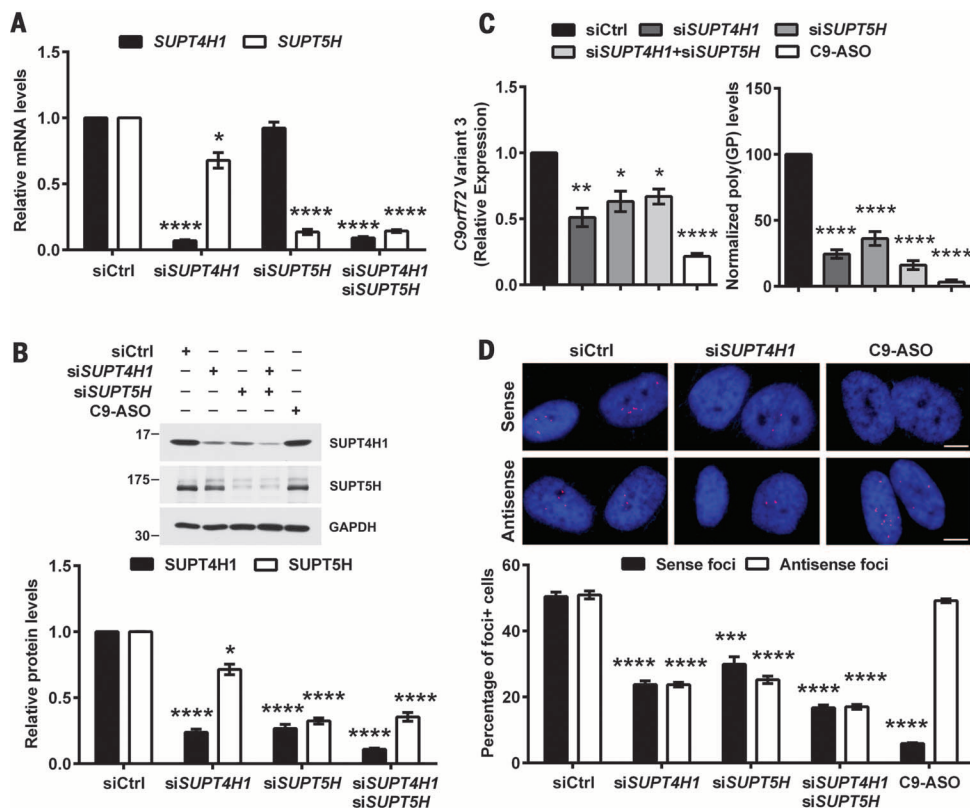


Fig. 3. Reduction of SUPT4H1 or SUPT5H in c9ALS fibroblasts inhibits production of *C9orf72* variant 3 mRNA, sense and antisense RNA foci, and DPR proteins. Cultured fibroblasts from three c9ALS patients were treated with a control siRNA (siCtrl), siRNA(s) directed against *SUPT4H1* and/or *SUPT5H* (siSUPT4H1 and siSUPT5H, respectively), or a *C9orf72*-repeat targeting ASO (C9-ASO) for 10 days. After treatment, mRNA levels of *SUPT4H1* and *SUPT5H* were determined by qPCR (A), and SUPT4H1 and SUPT5H protein expression was examined by immunoblot followed by densitometric quantification (B). (C) The effect of SUPT4H1 and/or SUPT5H depletion on *C9orf72* variant 3 mRNA expression and poly(GP) protein levels was examined by qPCR and immunoassay, respectively. (D) RNA FISH with probes for GGGGCC or GGCCCC RNA was used to detect foci containing sense or antisense repeats (red) in the nucleus of cells (blue, Hoechst 33258). The percentage of cells containing foci was then determined. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ as assessed by one-way ANOVA followed by Tukey's post-hoc analysis. Scale bar: 5 μ m.

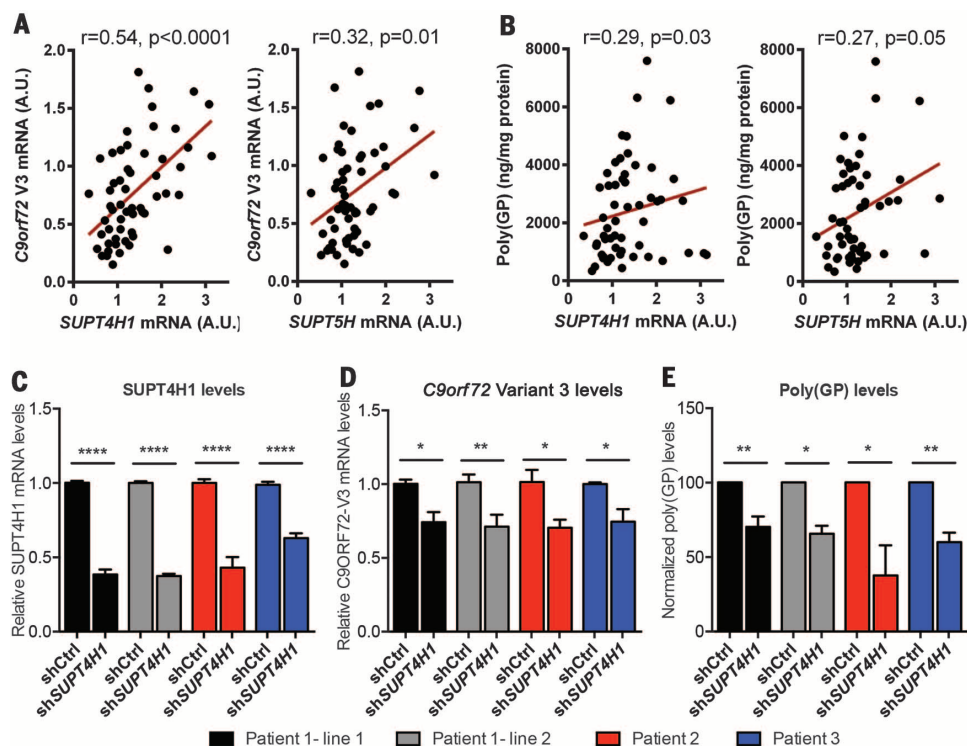


Fig. 4. Pathological features of c9FTD/ALS correlate with SUPT4H1 and SUPT5H expression in patient brain tissue, and partial knockdown of SUPT4H1 in *C9orf72* iPSC-derived cortical neurons reduces production of *C9orf72* variant 3 mRNA and DPR proteins. (A and B) SUPT4H1 and SUPT5H mRNA levels in cerebellar tissue from c9FTD/ALS patients were measured by qPCR and found to associate with cerebellar levels of *C9orf72* variant 3 mRNA measured by qPCR (A) and with poly(GP) proteins measured by immunoassay (B) using a Spearman's test of correlation. (C to E) Cortical neurons differentiated from each *C9orf72* iPSC line were aged for 8 weeks and then transduced with lentivirus expressing either a noncoding control shRNA or an shRNA against *SUPT4H1* mRNA for 6 days. Four iPSC lines derived from three patients with *C9orf72* repeat expansions were used. Knockdown of *SUPT4H1* mRNA (C) and *C9orf72* variant 3 mRNA expression levels (D) were determined by qPCR. Poly(GP) protein levels were examined by immunoassay (E). * $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$ as assessed by two-tailed unpaired t test from three independent neuronal differentiations.

respectively) simultaneously decreased both *SUPT4H1* and *SUPT5H* mRNA and protein levels (Fig. 3, A and B). Exposing fibroblasts to si*SUPT4H1* alone significantly decreased *SUPT4H1* mRNA and protein expression and also caused statistically significant, albeit less drastic, decreases in *SUPT5H* abundance (Fig. 3, A and B). Upon treatment of c9ALS fibroblasts with si*SUPT5H* alone, we observed the predicted reductions in *SUPT5H* mRNA and protein. Depletion of *SUPT5H* also led to a significant decrease in *SUPT4H1* protein abundance, although it had little effect on *SUPT4H1* mRNA (Fig. 3, A and B).

Reduction of *SUPT4H1* and/or *SUPT5H* abundance significantly decreased levels of *C9orf72* variant 3 mRNA and poly(GP) proteins, as well as foci formed of sense or antisense repeat RNA in c9ALS fibroblasts, but did not lead to any overt toxicity (Fig. 3, C and D, and figs. S6 and S7). In contrast, treatment of c9ALS fibroblasts with an ASO targeting the *C9orf72* sense transcript mitigated most of these features, with the key exception that foci formed of antisense GGCCCC-containing transcripts remained unaffected (Fig. 3, C and D). Thus, lowering the abundance of a single gene product, *SUPT4H1* or *SUPT5H*, decreased all three of the major pathological features of c9FTD/ALS: sense RNA foci, antisense RNA foci, and DPR proteins. Intriguingly, *SUPT4H1* and *SUPT5H* mRNA expression levels positively correlated with levels of *C9orf72* variant 3 mRNA (Fig. 4A) or poly(GP) proteins (Fig. 4B) in the cerebellum of human *C9orf72* repeat expansion carriers. Partial reduction in *SUPT4H1* abundance by lentiviral short hairpin RNA (shRNA) in 8-week-old cortical neurons differentiated from four independent lines of human-induced pluripotent stem cells from c9FTD/ALS patients (13, 15) (Fig. 4, C to E) resulted in decreased levels of *C9orf72* variant 3 mRNA (Fig. 4D) and poly(GP) proteins (Fig. 4E). To address whether there were global transcriptional changes upon depletion of *SUPT4H1*, we performed RNA sequencing on two

human fibroblast lines treated with *SUPT4H1* siRNA. As predicted by RNA-sequencing experiments in yeast and mouse (21), we found only a small subset of human genes that were differentially expressed after *SUPT4H1* knockdown (fig. S8), indicating that knockdown of *SUPT4H1* affects expanded repeat transcripts but has a minimal impact on other transcripts.

Here, we found that the transcription factor Spt4/*SUPT4H1* is required for the specific expression of hexanucleotide repeat expansions in the *C9orf72* gene. Previous studies have shown that injection of ASOs targeting *SUPT4H1* expression or genetic reduction in *Supt4h* levels in mouse models of Huntington's disease resulted in a mutant allele-specific reduction in huntingtin mRNA and protein abundance (27). This treatment reduced the amount of polyQ aggregates, delayed motor impairment, and prolonged survival (27). Our data suggest that similar strategies aimed at diminishing the function of *SUPT4H1*, such as ASOs or small molecules, could be pursued as an effective therapy for c9FTD/ALS.

REFERENCES AND NOTES

1. M. DeJesus-Hernandez et al., *Neuron* **72**, 245–256 (2011).
2. A. E. Renton et al., *Neuron* **72**, 257–268 (2011).
3. S. C. Ling, M. Polymenidou, D. W. Cleveland, *Neuron* **79**, 416–438 (2013).
4. T. F. Gendron, V. V. Belzil, Y. J. Zhang, L. Petrucelli, *Acta Neuropathol.* **127**, 359–376 (2014).
5. C. J. Donnelly et al., *Neuron* **80**, 415–428 (2013).
6. Y. B. Lee et al., *Cell Reports* **5**, 1178–1186 (2013).
7. A. R. Haeusler et al., *Nature* **507**, 195–200 (2014).
8. K. Mori et al., *Science* **339**, 1335–1338 (2013).
9. P. E. Ash et al., *Neuron* **77**, 639–646 (2013).
10. T. Zu et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4968–E4977 (2013).
11. T. F. Gendron et al., *Acta Neuropathol.* **130**, 559–573 (2015).
12. M. van Blitterswijk et al., *Acta Neuropathol.* **130**, 863–876 (2015).
13. S. Almeida et al., *Acta Neuropathol.* **126**, 385–399 (2013).
14. C. Lagier-Tourenne et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4530–E4539 (2013).
15. D. Sareen et al., *Sci. Transl. Med.* **5**, 208ra149 (2013).
16. T. F. Gendron et al., *Acta Neuropathol.* **126**, 829–844 (2013).
17. A. Hirtreiter et al., *Nucleic Acids Res.* **38**, 4040–4051 (2010).
18. P. B. Mason, K. Struhl, *Mol. Cell* **17**, 831–840 (2005).
19. A. G. Rondón, M. García-Rubio, S. González-Barrera, A. Aguilera, *EMBO J.* **22**, 612–620 (2003).
20. T. Wada et al., *Genes Dev.* **12**, 343–356 (1998).
21. C. R. Liu et al., *Cell* **148**, 690–701 (2012).
22. Y. Xue et al., *Cell* **152**, 82–96 (2013).
23. S. Mizielinska et al., *Science* **345**, 1192–1194 (2014).
24. B. D. Freibaum et al., *Nature* **525**, 129–133 (2015).
25. J. Chew et al., *Science* **348**, 1151–1154 (2015).
26. H. Tran et al., *Neuron* **87**, 1207–1214 (2015).
27. H. M. Cheng et al., *PLOS Genet.* **11**, e1005043 (2015).

ACKNOWLEDGMENTS

We thank D. Cerza and Y. Zu for technical support. This work was supported by NIH grants R01NS065317 (A.D.G.), R01NS09386501 (A.D.G., L.P.), R01NS073660 (A.D.G., N.M.B.), R01AG026251 (L.P.), R21NS079807 (Y.-J.Z.), R21NS094489 (C.N.C.), R01NS063964 (L.P.), R01NS077402 (L.P.), R21NS084528 (L.P.), P01NS084974 (L.P.), R01NS088689 (L.P.), R01ES020395 (L.P.), R01NS085812 (S.N.C. and T.-H.C.), R01NS079725 (F.-B.G.), Mayo Clinic Foundation (L.P.), Alzheimer's Association [NIRP-14-304425 (Y.-J.Z.)], Amyotrophic Lateral Sclerosis Association (Y.-J.Z., T.F.G., M.P., L.P., F.-B.G.), the National Science Foundation Graduate Research Fellowship (N.J.K.), the Robert Packard Center for ALS Research at Johns Hopkins (L.P., A.D.G., F.-B.G.), Target ALS (L.P., A.D.G., N.B., F.-B.G.), Association for Frontotemporal Degeneration (S.A.), Stanford University (S.N.C.), and the Glenn Foundation (A.D.G., N.M.B.). The induced pluripotent stem cell (iPSC) lines are available from the Gao laboratory under a material transfer agreement with the University of Massachusetts Medical School. R.C.P. has paid consulting relationships with Hoffmann–La Roche, Inc.; Merck, Inc.; Genentech, Inc.; Biogen, Inc.; and Eli Lilly & Co. RNA sequencing data are deposited with the National Center for Biotechnology Information Gene Expression Omnibus (GSE83484).

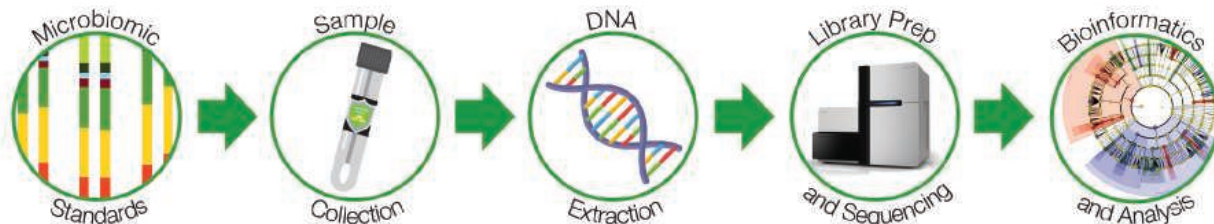
SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6300/708/suppl/DC1
Materials and Methods
Figs. S1 to S8
Table S1
References (28–40)
Data File S1

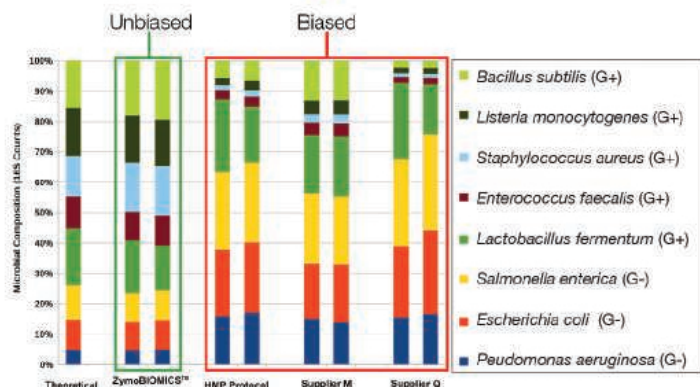
29 March 2016; accepted 6 July 2016
10.1126/science.aaf7791

ZymoBIOMICS™

A Comprehensive Solution for Microbiomics and Metagenomics



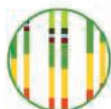
Standardizing Microbiomics



Accurate & Unbiased DNA Isolation

ZymoBIOMICS™ Microbial Community Standard was used to compare different DNA extraction protocols. DNA samples were profiled by 16S rRNA gene targeted sequencing. ZymoBIOMICS™ DNA Mini Kit yielded an unbiased community profile due to its efficient and robust lysis technology.

A Comprehensive Solution for Microbiomics



Optimize and Validate Microbiomic Workflows with the ZymoBIOMICS™ Microbial Community Standards



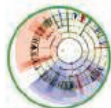
Stabilize Nucleic Acids at Collection using DNA/RNA Shield™ Collection Devices



Validated, nonbiased DNA Isolation using the ZymoBIOMICS™ DNA Isolation Kits



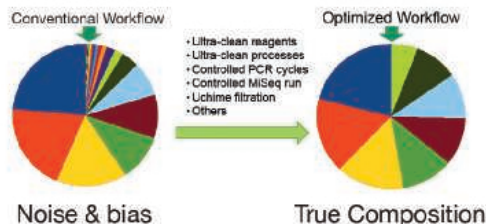
Ultra-clean, Low-bioburden reagents and services



Comprehensive, validated 16S rRNA and shotgun sequencing services for nonbiased, customizable, publication-quality data

Accurate & Reproducible Community Profiling

Validate and Optimize your Microbiomic Analyses using the ZymoBIOMICS™ Microbial Standards



The ZymoBIOMICS™ Microbial Community Standards are a well-defined and characterized community of Gram-positive bacteria, Gram-negative bacteria and yeast which are ideal for validating and optimizing microbiomic and metagenomic workflows from extraction to analysis. The pie chart on the left shows the genus-level microbial composition of the standard sequenced and analyzed using a conventional 16S rRNA gene sequencing workflow. The conventional workflow (left) prior to optimization indicates the presence of 43 bacterial genera, caused by artifacts like PCR chimera and process contamination. The optimized workflow (right) correctly shows the presence of 8 bacterial genera, which agrees with the actual composition of the standard.

Learn more today at www.zymoresearch.com/zymbiomics



AAAS 2017 ANNUAL MEETING

FEBRUARY 16-20 • BOSTON

SERVING SOCIETY THROUGH SCIENCE POLICY

Advance Registration Now Open

The 2017 meeting will focus on how to inform policies with the best available scientific evidence.

Register for the meeting and reserve a room now to receive lower rates. Join thousands of leading scientists, educators, policymakers, and journalists in Boston to discuss recent developments in science, technology, and policy.

Advance registration is available online.

aaas.org/meetings



SCIENCE & DIPLOMACY

SCIENCE & DIPLOMACY provides an open access forum for rigorous thought, analysis, and insight to serve stakeholders who develop, implement, or teach all aspects of science and diplomacy. Learn more about the latest ideas in science diplomacy and receive regular updates by following @SciDip on Twitter, liking the quarterly's page on Facebook (www.facebook.com/sciencediplomacy), and registering for free at www.sciencediplomacy.org/user/register.



WWW.SCIENCEDIPLOMACY.ORG

Science & Diplomacy is published by the Center for Science Diplomacy of the American Association for the Advancement of Science (AAAS), the world's largest general scientific society.

SCIENCE &
DIPLOMACY



The Industry's Preeminent Event
on Novel Drug Targets

14th Annual **Discovery** on **TARGET**

SEPTEMBER 19-22, 2016 | BOSTON, MA
WESTIN BOSTON WATERFRONT

CONFERENCE PROGRAMS

SEPTEMBER 20-21

- Targeting Histone Methyltransferases and Demethylases
- Targeting the Ubiquitin Proteasome System
- Targeting the Microbiome - Part 1
- GPCR-Based Drug Discovery - Part 1
- Advances in Gene Editing and Gene Silencing - Part 1
- Gene Therapy Breakthroughs
- Antibodies Against Membrane Protein Targets - Part 1
- Targeting Cardio-Metabolic Diseases
- Targeting Ocular Disorders

SEPTEMBER 21-22

- Targeting Epigenetic Readers and Chromatin Remodelers
- Kinase Inhibitor Discovery
- Targeting the Microbiome - Part 2
- GPCR-Based Drug Discovery - Part 2
- Advances in Gene Editing and Gene Silencing - Part 2
- Translating Cancer Genomics
- Antibodies Against Membrane Protein Targets - Part 2
- Metabolomics in Drug Discovery

SEPTEMBER 21-22 TRAINING SEMINAR: Data Visualization

SYMPOSIA SEPTEMBER 19

- Next-Generation Histone Deacetylase Inhibitors
- Strategies for Tackling Rare Genetic Diseases
- Understanding CRISPR: Mechanisms and Applications
- Autoimmunity – Small Molecule Approaches
- NK Cell-Based Cancer Immunotherapy

REGISTER TODAY!

Reference key code **L35** when registering

DiscoveryOnTarget.com

myIDP: A career plan customized for you, by you.



For your career in science, there's only one **Science**



**Recommended by
leading professional
societies and the NIH**

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.
- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- Ability to select which portion of your IDP you wish to share with advisors, mentors, or others.
- A certificate of completion for users that finish myIDP.

Visit the website and start planning today!

myIDP.sciencecareers.org

ScienceCareers In partnership with:





FT-NIR Spectrometer

For small business owners, technicians, laboratory managers, and method developers, a new Fourier transform-near infrared (FT-NIR) spectrometer is designed to bring rapid and reliable analysis for quality control and material verification. The Thermo Scientific Nicolet iS5N FT-NIR spectrometer uses the same proprietary software platform found across the entire Nicolet FTIR spectroscopy family. The OMNIC Lite software supports routine system operation, and the TQ Analyst EZ or Pro software can be used for calibration model development. Users manage and customize workflows within the intuitive graphic user interface provided by the OMNIC software suite, which includes Macros Basic and Macros Pro. FT-NIR is a workhorse technique for materials identification and quality control. A reliable, cost-effective material analysis tool designed to meet the needs of busy labs, the Nicolet iS5N is set to extend NIR technology and proven IR analysis into new arenas.

Thermo Fisher Scientific

For info: 800-955-6288

www.thermofisher.com/iS5N

ddPCR Multiplex Mutation Screening Kit

Five Droplet Digital PCR (ddPCR)

Multiplex Mutation Screening Kits are available (KRAS Q61, NRAS G12, NRAS G12/G13, NRAS Q61, and BRAF V600) for detecting several key actionable cancer mutations and the wild-type allele in a single reaction. The combination of ddPCR and multiplexing technologies increases the throughput of genetic analysis and allows clinical researchers to connect specific genetic alterations with diagnosis, treatment efficacy, and disease recurrence. A test-to-result turnaround time can be seen within 72 hours. Sample partitioning in ddPCR enables the detection of rare nucleic acid sequences in a complex background of competing sequences, while minimizing the effect of inhibitory substances.

The kits identify mutations present at allelic frequencies at or below 0.5%, making them ideal for use with FFPE and liquid biopsy samples. Their precision and sensitivity also enable early detection of important biomarkers.

Bio-Rad

For info: 510-741-1000

www.bio-rad.com



Confocal Laser Scanning Microscope

Zeiss introduces the LSM 800 confocal laser scanning microscope for materials applications and analysis in research and industry. The LSM 800 enables precise, 3D imaging of microstructures and surfaces. The combination of confocal fluorescence and other contrasting techniques [brightfield, darkfield, polarization, and Zeiss's unique Circular Differential Interference Contrast (C-DIC)] in one instrument enables high-precision examination of nanomaterials, metals, polymers, and semiconductors with maximum information content. The system offers precise capture of 3D topography and the investigation of nanometer-scale structures without causing surface damage. The combination of the LSM 800 laser scanning module with Zeiss's Axio Imager enables a wide variety of configurations in hardware, for example, objectives, stages, and illumination. The LSM 800 comes with the newest version of Zeiss Efficient Navigation (ZEN) imaging software, which includes an open application development interface for data exchange with established third-party analysis and research software.

Zeiss

For info: +49-(0)-3641-64-3949

www.zeiss.com

Bioinformatics Software

The BaseSpace Informatics Suite accelerates genomic data analysis for sequence labs. It is the first cloud-based, sample-to-answer informatics infrastructure to enable precision medicine and genomics research. The integrated informatics solution enables users to engage with a unified, comprehensive platform, from laboratory information management, analysis, and sharing to sample interpretation, reporting, and advanced cohort analytics. The BaseSpace Suite reduces the time from sample submission to answer. It improves the overall accuracy of results without the need to integrate point solutions from multiple vendors, and is ideal for numerous labs, including clinical and translational labs and research institutions. The BaseSpace Suite is built to be cloud-ready, providing a secure, scalable solution, simplifying data storage issues, and providing a highly collaborative environment. The system provides preconfigured and customizable workflows and analysis pipelines. The platform is intuitive and user-friendly for a myriad of clients—from the clinical scientist to the lab director and researcher.

Illumina

For info: 858-202-4500

www.illumina.com

Carbon/Sulfur Analyzers

The EMIA-Pro and Expert carbon/sulfur analyzers offer nondispersive infrared (NDIR) measurement capabilities, optimizing their range from 1.6 ppm to 6.0% for carbon, and 2 ppm to 1.0% for sulfur. Equipped with a unique carbon monoxide detector, the analyzers can be used for a wide variety of inorganic materials such as steel, cokes, catalysts, nonferrous alloys like aluminum, and lithium-ion battery materials. A new cleaning mechanism allows increased cleaning efficiency compared to that of conventional carbon/sulfur analyzers. This mechanism enables users to repeat their measurements up to 200 times, with less time spent for maintenance and cleaning. A new, robust design with fewer components enables more reliable use and less service

requirements. The EMIA-Pro and Expert offer increased measurement efficiency, with measurement, display of results, and cleaning taking about 70 seconds. They are also equipped with a navigation system that recommends the most suitable conditions for samples, troubleshoots errors, and alerts users to perform maintenance.

Horiba Scientific

For info: 732-494-8660

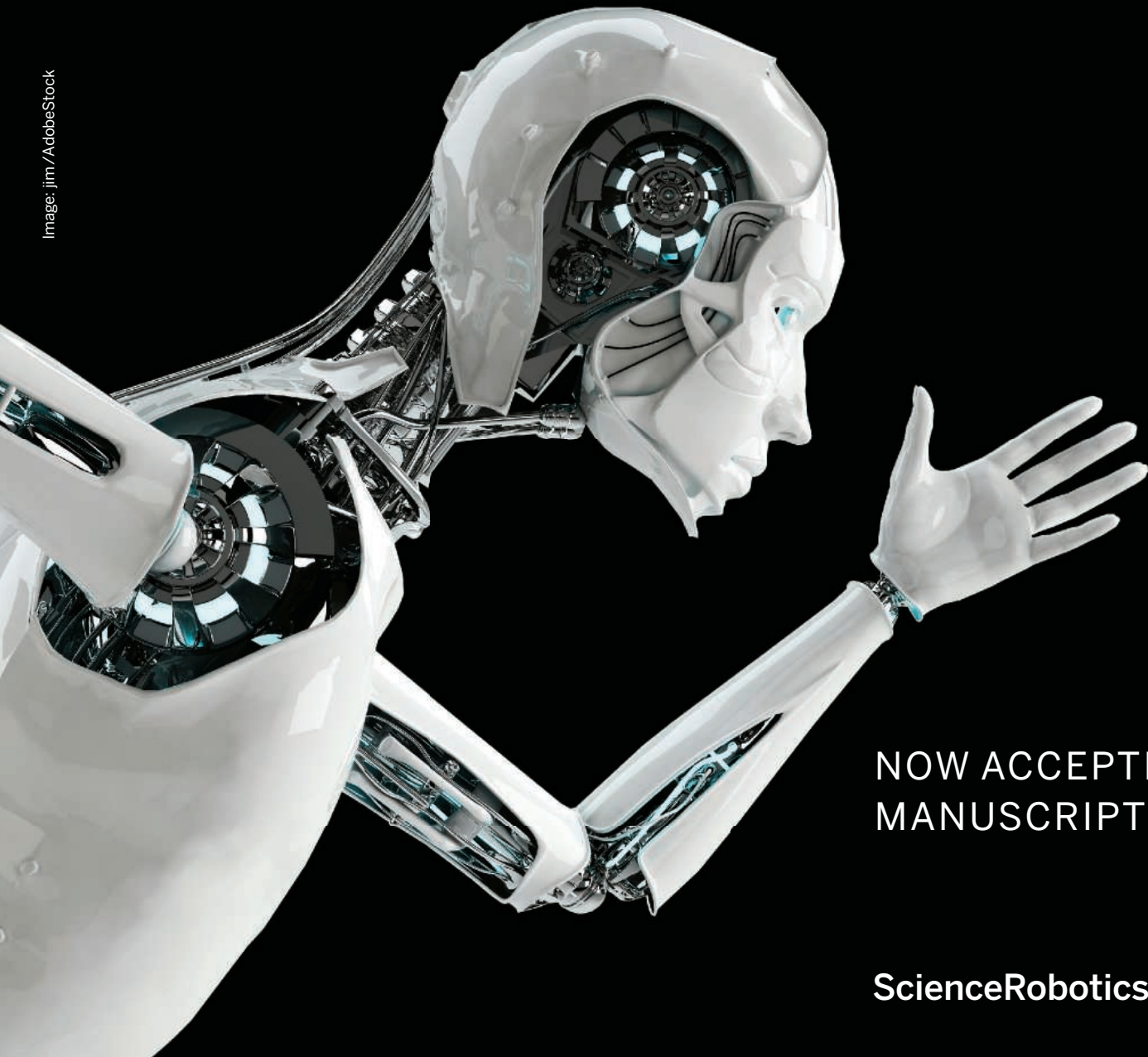
www.horiba.com/scientific

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

Be Among the First to Publish in ***Science Robotics***

Image: jim / AdobeStock



NOW ACCEPTING
MANUSCRIPTS

ScienceRobotics.org

Science Robotics is a unique journal created to help advance the research and development of robotics for all environments. *Science Robotics* will provide a much-needed central forum to share the latest technological discoveries and to discuss the field's critical issues.

Join in the excitement for the Fall 2016 debut!

ScienceRobotics
AAAS

Accelerate the joy of your discovery

Add Aldrich™

KitAlysis™ High-Throughput Screening Kits provide the optimal solution to quickly and efficiently identify suitable catalytic cross-coupling reaction conditions.

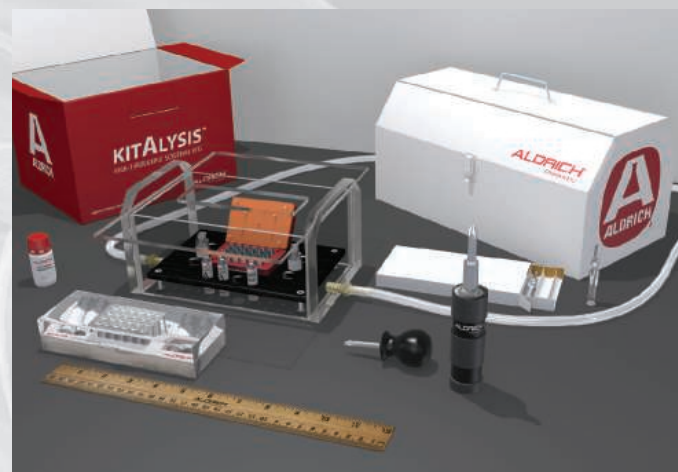
This unique off-the-shelf microscale screening platform:

- Requires only ~100 mg total of each substrate to run 24 distinct cross-coupling reactions
- Includes all chemical components
- Eliminates the need to set up reactions in a glove box with KitAlysis labware (sold separately)
- Provides comprehensive online tools including user guides, tutorial video and reaction calculators
- Has been designed and validated by practicing medicinal and process chemists in some of the largest pharma companies in the world

Find out how KitAlysis can accelerate your research, visit

Aldrich.com/KitAlysis

Questions? Contact **KitAlysis@sial.com**



Cross-coupling screening kits include:



Four sets of preweighed catalysts in glass microvials with stir bars.



Ampulized Sure/Seal™ anhydrous solvents and liquid base solutions.



Prewieghed, finely ground bases. Substrate vials and stir bars.



Internal standard for ease of analysis.

By Jeffrey J. McDonnell

The 1-hour workday

When I was an assistant professor, I felt constantly overwhelmed. I had classes to teach, relationships with new colleagues to navigate, a lab group to assemble, and an infant at home—not to mention research to conduct and papers to publish. To get ahead, I took on any opportunities that were offered, including membership on various editorial boards and professional committees. Despite working like a madman, my productivity as measured by paper output was meager. I simply could not find time in my day for undistracted writing. And when I did find the time after an extended stretch away from writing, the warm-up period to get back into the paper was often long, further slowing my progress.

At first I thought that this kind of frustration was routine for academic researchers. But as the years passed, I noticed a few senior colleagues who published with impressive regularity and always had a paper in the works. When I asked them what their secret was, I found that they prioritized doing small amounts of focused writing every day. I've since developed my own version of this approach. I call it the 1-hour workday, referring to the short, sacrosanct period when I do what I see as the "real" work of academia: writing papers.

First thing in the morning is when I'm at my mental best, and when I'm still most in control of my time, so I now use the first hour of my day to write. For me, it's best done from home. I've developed something of a ritual: I wake up early, make an espresso, and write until I'm spent—or until distractions like email or the day's deadlines and meetings start to intrude. This is usually about an hour, some days a little less and some days more. I've found that, like hitting a ball in golf, regular writing is easier if I tee it up. I plan my early morning writing the night before. It is in my calendar and on my to-do list, with details about which paper and section I will be working on.

This routine has transformed my work life. Instead of the frustration that frequently plagued me early in my career, now—no matter how work proceeds after I've completed my writing time—I go home at the end of the day with the satisfaction of having accomplished something.

I have in no way mastered the writing game, but my 1-hour workday has certainly increased my academic output. And by keeping me focused and in practice, it has improved the quality of my writing and made the process



***"I wake up early,
make an espresso, and
write until I'm spent."***

much more enjoyable. It also offers an opportunity for deep thinking. I remember rarely having any such thinking time when I started out as a professor, but now I find that my daily keystrokes can lead to new ideas. When I string together days of successful writing, ideas flow and new connections present themselves even when I'm away from my keyboard, particularly on my bike ride to work or when I'm reading for pleasure.

Many days, "writing" means editing and revising the work of others. This work can sometimes be a slog, but I keep my spirits up by thinking of it as a game of table tennis. My goal is to return the serve—when done quickly, this greatly improves the game.

In any sport, one must stay toned and conditioned. If I fall out of practice, I quickly lose that fitness, and my writing and editing become labored. So, even though distractions abound, I protect that daily work-out at the keyboard during the first precious work hour of the day. I've learned that writing does not need long stretches of uninterrupted time. Focus and regularity are what matter. I now advise my Ph.D. students and postdocs who are going on to faculty positions to adopt daily writing as an early-career habit so that they don't repeat my years of writing frustration. At any career stage, a daily writing ritual can help improve performance—and, perhaps most importantly, job satisfaction. ■

Jeffrey J. McDonnell is a professor in the School of Environment and Sustainability at the University of Saskatchewan in Saskatoon, Canada, and Sixth Century Chair at the University of Aberdeen in the United Kingdom. He thanks his Ph.D. students and postdocs for feedback.



There's only one **Science**

Science Careers Advertising

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.

Tracy Holmes
Worldwide Associate Director
Science Careers
Phone: +44 (0) 1223 326525

THE AMERICAS

E-mail: advertise@sciencecareers.org
Fax: +1 (202) 289 6742

Tina Burks
Phone: +1 (202) 326 6577

Nancy Toema
Phone: +1 (202) 326 6578

Online Job Posting Questions
Phone: +1 (202) 312 6375

EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD

E-mail: ads@science-int.co.uk
Fax: +44 (0) 1223 326532

Sarah Lelarge
Phone: +44 (0) 1223 326527

Kelly Grace
Phone: +44 (0) 1223 326528

Online Job Posting Questions
Phone: +44 (0) 1223 326528

JAPAN

Katsuyoshi Fukamizu (Tokyo)
E-mail: kfukamizu@aaas.org
Phone: +81 3 3219 5777

Hiroyuki Mashiki (Kyoto)
E-mail: hmashiki@aaas.org
Phone: +81 75 823 1109

CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND

Ruolei Wu
E-mail: rwu@aaas.org
Phone: +86 186 0082 9345

Danny Zhao
E-mail: dzhao@aaas.org
Phone: +86 131 4114 0012

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. *Science* reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. *Science* encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

ScienceCareers.org



**Learn more and
keep your job search
out of the cheap seats.**

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources
on **ScienceCareers.org**.

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS



Why choose this Postdoc Feature your advertisement?

- Relevant ads lead off the career section with a special "Postdoc" banner.

Expand your exposure by posting your print ad online:

- Link on the job board homepage directly to postdoc positions
- Dedicated landing page for postdoc positions.

Produced by the *Science*/AAAS Custom Publishing Office.



SCIENCECAREERS.ORG

ScienceCareers

FROM THE JOURNAL *SCIENCE* AAAS

To book your ad: advertise@sciencecareers.org

The Americas: 202 326 6582

Europe/RoW: +44(0) 1223 326500

China/Korea/Singapore/Taiwan: +86 186 0082 9345

Japan: +81 3 3219 5777

Why not change the world?

Tenure-Track Position Open in Biological Sciences

The Department of Biological Sciences at Rensselaer Polytechnic Institute in Troy, NY invites applications for a tenure-track faculty position at the rank of Assistant Professor. Applicants are expected to establish a vigorous and externally funded research program in any area of Cell Biology, Developmental Biology, or Neurobiology.

Candidates must have an outstanding record of scientific achievement demonstrated by publications in peer-reviewed journals. At minimum, applicants must hold a Ph.D. degree or foreign degree equivalent with postdoctoral experience in a relevant field. The candidate will be expected to teach at both the undergraduate and graduate levels in biology, biochemistry, biophysics, bioinformatics, neurobiology, or closely related disciplines. The successful applicant will be housed in the Center for Biotechnology and Interdisciplinary Studies (<http://biotech.rpi.edu/>), which boasts an impressive array of state-of-the-art facilities (<http://biotech.rpi.edu/facilities>).

The Department of Biological Sciences at Rensselaer (<http://science.rpi.edu/biology>) is highly interdisciplinary and the core of the life sciences at Rensselaer. We are a broad-based interactive community of faculty, students, and staff whose research interests encompass nearly every area of modern biology including cell and developmental biology, macromolecular structure and function, computational biology and bioinformatics, biochemistry and bioenergetics, biophysics and single molecule mechanics, neuroscience, circadian clock networks and biological outputs, microbiology and host-pathogen interactions, freshwater ecology, nanobiology and biotechnology, and synthetic biology. The Biological Sciences faculty have strong collaborations with colleagues within the Center for Biotechnology and Interdisciplinary Studies and other departments across campus.

Rensselaer is committed to academic excellence and diversity within the faculty. The University is located in Troy, NY, in the Capital District, which hosts several other Universities and Colleges and the State Government of New York.

Successful candidates will have duties that include teaching at the undergraduate and graduate level in the department, development and maintenance of robust programs of research and scholarship, as well as service to the department, the School of Science, and Rensselaer.

Applications must be sent as a single PDF document containing a cover letter, curriculum vitae, statement of research accomplishments and goals (~3 pages), and a one-page description of teaching interests and philosophy to biology-chair@rpi.edu. Applicants must also arrange for the submission of three letters of reference to biology-chair@rpi.edu. Questions about this position may be directed to Lee Ligon, Search Committee Chair, at ligonl@rpi.edu or Susan Gilbert, Department Head, at sgilbert@rpi.edu.

Application review will begin on September 1, 2016 and continue until the position is filled. The appointment will begin August 2017.

Rensselaer

We welcome candidates who will bring diverse cultural, ethnic, and national and international perspectives to Rensselaer's work and campus communities. Rensselaer Polytechnic Institute is an Affirmative Action/Equal Opportunity, Race/Gender/Veterans/Disability Employer.



Faculty Careers

Issue date: September 16

Book ad by August 30

Ads accepted until Sept 9 if space allows

Issue date: October 7

Book ad by September 20

Ads accepted until Sept 30 if space allows

For recruitment in science, there's only one *Science*.

Hiring Faculty? Whatever your timing, we've got two special features for your faculty ads this fall! The September 16 feature offers advice on how to develop skills for reviewing grants and papers. The October 7 feature covers business principles for researchers. Reach *Science* readers and share opportunities at your university.

What makes *Science* the best choice for recruiting?

- Read and respected by 400,000 readers around the globe
- 62% of our weekly readers work in academia and 65% are Ph.D.s. *Science* connects you with more scientists in academia
- Your ad dollars support AAAS and its programs, which strengthens the global scientific community.

Why choose these Faculty Features for your advertisement?

- Relevant ads lead off these career sections with a special Faculty banner
- October 7 issue will be distributed at the American Society of Human Genetics meeting, 18–22 October, Vancouver.

Expand your exposure by posting your print ad online:

- Link on the job board homepage directly to Faculty jobs
- Dedicated landing page for faculty positions.

Deliver your message to a global audience of targeted, qualified scientists.

129,574

subscribers in print every week

352,966

monthly unique browsers on ScienceCareers.org

65 %

of our weekly readers are Ph.D.s



Produced by the *Science*/AAAS Custom Publishing Office.

SCIENCECAREERS.ORG

Science Careers

FROM THE JOURNAL SCIENCE AAAS

To book your ad: advertise@sciencecareers.org

The Americas
+202 326 6582
Japan
+81 3 3219 5777

Europe/RoW
+44 (0) 1223 326500
China/Korea/Singapore/Taiwan
+86 186 0082 9345

POSITIONS OPEN

MICROBIOLOGIST

The Department of Biological Sciences, California State University, East Bay, seeks a microbiologist for appointment to a tenure-track position at the level of **ASSISTANT PROFESSOR** beginning Fall 2017. First consideration will be given to individuals with a background in microbiome research, molecular mechanisms of microbial pathogenesis, virulence and/or quorum sensing, or host-pathogen interactions. We seek applicants who are committed to excellence in teaching and the development of a student-centered research program in a culturally diverse intellectual community. The primary teaching responsibility will include undergraduate and Master's level courses in microbiology. Additionally, applicants will be expected to establish and maintain an externally funded research program. A Ph.D. and postdoctoral experience are required. Submit cover letter, curriculum vitae, separate teaching, research, and diversity statements, recent publications most pertinent to proposed research program (three maximum), and teaching evaluations if available. Submit application materials, and arrange to have three letters of letters of recommendation, uploaded directly to Interfolio (website: <http://www.interfolio.com>). For full consideration, applications must be received by October 7, 2016. Direct inquiries to e-mail: microsearchcommittee@csueastbay.edu. *Equal Opportunity Employer.*

PRINCETON UNIVERSITY DEPARTMENT OF CHEMISTRY ASSISTANT PROFESSOR

The Department of Chemistry at Princeton University invites applications for a **TENURE-TRACK ASSISTANT PROFESSOR** position in all areas of chemistry. We seek a faculty member who will create a climate that embraces excellence and diversity with a strong commitment to research and teaching that will enhance the work of the department and attract and retain a diverse student body. We strongly encourage applications from members of all underrepresented groups. Candidates are expected to have completed the Ph.D. in chemistry or a related field at the time of appointment. Applicants should submit a description of research interests, curriculum vitae, a list of publications, and contact information for three referees online at website jobs.princeton.edu/applicants/Central?quickFind=67134. The deadline for applications is October 15, 2015. *Princeton University is an Equal Opportunity Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, or any other characteristic protected by law. This position is subject to the University's background check policy.*

**Post your jobs
Fast and Easy**



ScienceCareers

employers.sciencecareers.org



Assistant, Associate or Professor in Biochemistry & Molecular Genetics

In a major expansion, the School of Medicine at the University of Virginia (www.virginia.edu/bmg) plans to hire junior tenure-track and established tenured faculty members (open rank). As part of this initiative, the Department of Biochemistry & Molecular Genetics seeks candidates who address contemporary problems in cancer or other diseases, regenerative medicine, aging, immunotherapy or the basic biology of the mammalian cell nucleus.

Candidates must hold a Ph.D. and/or M.D. in the life sciences or computational biology and must have demonstrated the capacity to develop and/or maintain a high-quality independent research program. Applicants with outstanding credentials from all fields will be considered. Support will include attractive start-up packages, laboratory space, and an exceptionally interactive research environment. Clinically qualified applicants will be considered for a joint appointment in the appropriate clinical department.

To apply, visit <https://jobs.virginia.edu> and search on posting number 0619017. Complete a candidate profile online, attach a cover letter, curriculum vitae, and statement of research interest (**must be two pages or less including references**).

Applicants for tenure-track positions should have references submit three letters of recommendation electronically, **directly to hn6u@virginia.edu** with the applicant's name in the subject heading. For senior applicants seeking tenured positions, letters are not required at the outset and short-listed candidates will be contacted before asking potential referees.

Positions will remain open to applicants until filled.

For additional information, please contact Helen Norfleet-Shiflett, Department of Biochemistry & Molecular Genetics, via telephone (434) 924-1667 or email hn6u@virginia.edu.

The University of Virginia is an equal opportunity/affirmative action employer. Women, minorities, veterans and persons with disabilities are strongly encouraged to apply.

**Advance your career
with expert advice from
Science Careers.**



Download Free Career Advice Booklets!

ScienceCareers.org/booklets

Featured Topics:

- Networking
- Industry or Academia
- Job Searching
- Non-Bench Careers
- And More



ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

